An *in vivo* study into the metabolic reprogramming of hepatocellular carcinoma

Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium

(Dr. rer. nat.)

im Fach Biologie

Eingereicht an der

Lebenswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

Olga Vvedenskaya

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Bernhard Grimm

Gutachter/innen: 1. Prof. Dr. Markus Landthaler

2. Prof. Dr. Christine Sers

3. Prof. Dr. Thorsten Cramer

Tag der mündlichen Prüfung: 15.09.2017

Forty-eight days ago I graduated from the faculty with honors, but honors is one thing and hernia is another.

-Mikhail Bulgakov, *A Country Doctor's Notebook, 1925-1926*

Сорок восемь дней тому назад я кончил факультет с отличием, но отличие само по себе, а грыжа сама по себе.

-Михаил Булгаков, *Записки юного врача, 1925-1926*

Summary

The present work evaluates the role of metabolism in development and progression of hepatocellular carcinoma (HCC). This study focuses on changes of central metabolic pathways, including glycolysis, gluconeogenesis, tricarboxylic acid (TCA) cycle and other processes involved in cellular metabolism and known to be dysregulated during cancer formation. Comprehensive multiomics analyses, such as proteomics, metabolomics and targeted genome sequencing, were applied in order to better understand HCC developmental mechanisms *in vivo*. Two main systems were studied: the ASV-B mouse model and clinical samples from human patients. The human cohort was composed of biopsy and surgery material from 95 patients: 47 HCC and 48 non-HCC.

Proteomic data from both mice and humans show a clear downregulation of the main energy-producing pathways in HCC. Glycogen metabolism, *de novo* glucose synthesis, glutamine uptake to the TCA cycle, approximately 60% of enzymes of TCA cycle, and transport of pyruvate to mitochondria are downregulated in HCC. An isoform switch at various levels of central carbon metabolism was demonstrated in this work. Both mice and humans with HCC reveal isoform switches at the level of phosphoglycerate mutases and pyruvate kinases. The key features are found in both mouse and human, showing a universal metabolic HCC fingerprint. This study also demonstrates that proteomic analysis of the bioptate material is a strong and sufficient molecular diagnostic tool for research in cancer: the proteomic analysis of liver material allows the distinction of tumor samples from non-tumor samples and also to track the level of disease progression.

Targeted genome sequencing revealed that no clear distinction between cancer and precancerous conditions could be made exclusively from the mutation analysis. Human metabolomic data remains inconclusive, possibly due to the different sources of tissue samples.

Zusammenfassung

Die vorliegende Arbeit untersucht die Rolle des Metabolismus in der Entstehung und Progression des Hepatozellulären Karzinoms (HZK). Der Schwerpunkt der Studie liegt auf Veränderungen zentraler Stoffwechselwege, unter anderem der Glykolyse, der Gluconeogenese, des Citratzyklus und anderer Prozesse des Zellstoffwechsels. Umfassende Multiomikanalysen, wie etwa Proteomik, Metabolomik und gezielte Genomsequenzierung wurden angewandt, um *in vivo* die Mechanismen der HZK Entstehung zu verstehen. Es wurden zwei Systeme untersucht: das ASV-B Mausmodell und klinische Patientenproben. Die Kohorte bestehend aus Biopsien und Resektaten von 95 Patienten umfasste 47 Fälle von HZK und 48 Fälle ohne HZK.

Das Proteom des Mausmodells und der Patientenkohorte zeigen eine deutliche Herabregulierung wesentlicher Energie bereitstellender Kreisläufe im HZK: Glykogenstoffwechsel, *de novo* Synthese von Glukose, Glutaminaufnahme in den Citratzyklus, des weiteren sind 60% der Enzyme des Citratzyklus, und des Transports von Pyruvat in Mitochondrien im HZK herabreguliert. In dieser Arbeit wurde ein Isoformenwechsel auf mehreren Ebenen des zentralen Kohlenstoffmetabolismus gezeigt. Sowohl das Mausmodell, als auch die Gewebeproben von HZK-Patienten weisen Isoformenwechsel der Phosphoglyzeratmutasen und der Pyruvatkinasen auf. Die Hauptmerkmale finden sich sowohl in Modellmäusen, als auch in Patienten, und stellen so einen universalen metabolomischen Fingerabdruck des HZK dar. Darüber hinaus demonstriert diese Studie, dass die Proteomanalyse von bioptischen Material ein aussagekräftiges und ausreichendes molekular-diagnostisches Instrument für die Krebsforschung ist: die Proteomanalyse von Lebermaterial erlaubt die Unterscheidung von Tumorgewebe und tumorfreien Proben und die Dokumentation des Krankheitsverlaufs.

Keywords

English: hepatocellular carcinoma, proteomics, metabolomics, central carbon metabolism, mutation, mass spectrometry.

German: hepatozelluläres Karzinom, Proteomik, Metabolomik, zentraler Kohlenstoffmetabolismus, Mutation, Massenspektrometer.

Outline

Hepatocellular carcinoma (HCC) is a malignant tumor which originates from hepatocytes and causes death in 95% of cases within five years from the established diagnosis. The deficit of fundamental knowledge about processes taking place in HCC development leads to a lack of diagnostic and treatment options, and therefore to the poor prognosis of the disease. In this work I aim to elucidate the role of metabolism in HCC development and progression. I am particularly interested in central metabolic pathways, e.g. glycolysis, gluconeogenesis, tricarboxylic acid (TCA) cycle and other processes involved in cellular metabolism known to be dysregulated during cancer formation. In this regard it is important not only to study metabolism and to investigate the features which can help differentiate HCC from non-HCC, but also to explore metabolic and genetic heterogeneity among the HCC samples. In this study I performed comprehensive multiomics analyses, which included proteome, metabolome and genetic background analyses, to better understand HCC developmental mechanisms in two main systems: the ASV-B mouse model and clinical samples from human patients. Defining the HCC signature changes in metabolism could shed light on universal HCC biomarker discovery.

The ASV-B mouse model is a transgenic model which shows dysplastic nodule formation at the age of 8 weeks. By the 12th week, adenomas in the liver can be observed, and differentiated HCC develops by the 16th week in these mice, the age at which they were used in this study. C57BL/6J mice were used as a control (Dubois, et al., 1991). The human cohort was composed of biopsy and surgery material from 95 patients: 47 HCC and 48 non-HCC (which comprised various other conditions affecting the liver such as fatty liver, hepatitis B or C, fibrosis and others). Several issues had to be considered during the cohort formation to reduce the great heterogeneity among patients. We excluded patients, according to their anamnesis, with exposure to certain hepatotoxins, as well as patients with long-term usage of certain medication, patients with known exposure to certain hepatotoxins, patients receiving lifelong medication, and those with hereditary conditions or immune-mediated diseases. Current hepatitis viral load on the day of material acquisition was considered, and separate groups of patients with hepatitis virus B or C were formed. Patients with other detected viruses (e.g. HIV[1]) were excluded from the cohort. To the best of our knowledge patients did not get any treatment prior to surgery or biopsy procedures.

[1] HIV – human immunodeficiency virus

To obtain data about proteins and metabolites in liver samples I used mass spectrometry (MS). For protein analysis, a shotgun proteomics approach was used, which combines nano Ultra-High Performance Liquid Chromatography (nanoUHPLC) with MS. For the analysis of metabolites, we used gas chromatography time-of-flight (GC-TOF) MS. This method allows a specific separation and quantification of central-carbon metabolism-related metabolites. The main cancer-related somatic mutations in human material were analyzed in a targeted way with Ion Personal Genome Machine (PGM) Torrent and an Ion AmpliSeq Cancer HotSpot Panel v2.

Proteomic data from the ASV-B mouse model showed a specific picture of metabolic reprogramming in the damaged liver. In cancerous liver tissue glycogen metabolism was impaired via the inhibition of glycogen synthase 2 (Gys2). Several enzymes of glycolysis showed isoform switches e.g. phosphoglycerate mutases (Pgm), pyruvate kinases (Pk)), which represents one of the most striking features of metabolic changes in cancer. Gluconeogenesis was downregulated at the level of phosphoenolpyruvate carboxykinase 1 (Pck1) and glucose-6-phosphatase catalytic-subunit (G6pc). Transport of pyruvate to mitochondria was inhibited at the level of mitochondrial pyruvate carriers 1 and 2 (Mpc 1, 2). 60% of TCA cycle enzymes were downregulated, for example citrate synthase (Cs) and succinate dehydrogenases A, B and C (Sdha, b, c). Enzymes, which are responsible for the glutamine uptake into the TCA cycle had decreased expression in the HCC samples: glutamate pyruvate transaminase (Gpt), glutamate pyruvate transaminase 2 (Gpt2), and glutaminase 2 (Gls2) were downregulated at the proteomic level. At the metabolic level the higher amounts of glucose-6-phosphate, fructose-6-phosphate and lactic acid were detected. Levels of TCA cycle metabolites were higher in HCC as well: fumaric and malic acids were higher in HCC compared to control.

In the proteomic profile of the human samples a similar signature of tumor-specific metabolic reprogramming was found. At the proteomic level, glycogen metabolism was downregulated via dysregulation of glycogen phosphorylase L (PYGL) and glycogen synthase 2 (GYS2). Glucose *de novo* synthesis, which is essential for an organism during the fasting time, was downregulated at the level of fructose-1,6-bisphosphatase 1 (FBP1): FBP1 expression was approximately 7 times less in HCC compared to control group. This observation suggests that liver cannot maintain one of its' main metabolic functions - gluconeogenesis. Glycolysis was characterized by an upregulation of some enzymes, such as hexokinases (HK), and by an isoform switch at the levels of phosphoglycerate mutases (PGM), aldolases (ALDO), enolases (ENO), pyruvate kinases (PK), and lactate dehydrogenases (LDH). 65% of the enzymes of the TCA cycle

were downregulated, and isocitrate dehydrogenase (IDH) showed a switch between subunits. Glutaminolysis was downregulated at the level of glutamate pyruvate transaminase (GPT) and glutaminase 2 (GLS2). Due to the tremendous heterogeneity of the phenotypes of human patients, metabolic data was mostly inconclusive, thus could not be used as a reliable source of data. This diversity may come from different environmental conditions, diets, and associated diseases of patients in cohort.

Targeted genome sequencing was used in order to track well characterized cancer-related somatic mutations. The same mutations were detected in samples with fatty liver, hepatitis B and C, fibrotic liver, cirrhosis, and HCC. No clear distinction between cancer and precancerous conditions could be made exclusively from the mutation analysis. However, several CTNNB1[1] mutations were specific for tumor samples and were detected in 15% of HCC. This observation can indicate that CTNNB1 mutation is not a necessary but sufficient factor for the CCM changes, and, when present, affects metabolism in a drastic way which has an impact on major liver functions. Thus, the above mentioned CTNNB1 mutations should further be validated on larger cohorts for potential use as a prognostic biomarker of survival or metastatic status.

Taken together, proteomic data from both mice and humans showed a clear dysregulation of the main energy-producing pathways in HCC. Glycogen metabolism (via PYGL and GYS2), *de novo* glucose synthesis (on various levels), glutamine uptake to the TCA cycle (via GPT, GLS2), and approximately 60% of enzymes of TCA cycle (SDHs, SUCLs, ACO etc.), and transport of pyruvate to mitochondria (MPCs) were downregulated in HCC. The proteomic data demonstrated one of the most common features of cancer cells: metabolic reprogramming. An isoform switch, which is one of the fundamental mechanisms in reprogramming, was demonstrated in this work. Both mice and humans with HCC revealed isoform switches at the level of phosphoglycerate mutases and pyruvate kinases. Despite the diversity of genetic background of human samples the proteomic HCC signature is very robust in both bioptate and surgery material. The key features were found in both mouse and human, showing a universal metabolic HCC fingerprint. This study also demonstrates that proteomic analysis of the bioptate material is a strong and sufficient molecular diagnostic tool for research in cancer: the proteomic analysis of liver material allows the distinction of tumor samples from non-tumor samples and also to track the level of disease progression.

[1] CTNNB1 - Catenin Beta 1 (β-catenin)

Acknowledgments

Making a parallel with the above mentioned quote by Mikhael Bulgakov, I would like to start over my story with separating two things. For me my work on Ph.D. thesis was one thing, and my personal odyssey was another thing. My journey of four and a half years into Ph.D. started in February 2013, when I joined Berlin School of Integrative Oncology and lab of Prof. Dr. Thorsten Cramer, to whom I am incredibly thankful for giving me a start in Berlin and Charite, and for the further work together on this project.

However, things change, so as my Ph.D. supervision had to be reconsidered one year later. The successful completion of this adventure would have been impossible without my supervisor, teacher and great friend Dr. Stefan Kempa. Stefan, I cannot express how I am thankful to you for giving me an astonishing example of how to science and how to live. You were always there when I needed you, even though I myself was not able to understand it. You gave me all the freedom to make my own decisions, and kept an eye on me when I was working on them. Thank you for always believing in me and letting me know that.

I strongly believe, that no one can do good science alone. This project complemented a lot from the help of our great collaborators: Prof. Dr. Christine Sers, Dr. Soulafa Mamlouk and Andrea Menne. I have to additionally point out, that Christine Sers was the first person I had an interview with during the BSIO application round in December 2012. The value of an inspiration she gave me that day is hard to overestimate.

I am also thankful to the members of my Ph.D. thesis committee who are kind enough to support me during my last steps of this journey: Prof. Dr. Andreas Herrmann, Prof. Dr. Markus Landthaler, and Dr. Benedikt Beckmann, and of course Prof. Dr. Christine Sers and Prof. Dr. Thorsten Cramer.

My amazing lab members were always there to support, encourage, teach and help me. Birte Arlt, Dr. Chris Bielow, Alina Eisenberger, Martin Forbes, Fardad Ramezani, Dr. Raphaela Fritsche, Jenny Grobe, Dr. Henning Kuich, Dr. Guido Mastrobuoni, Dr. Tobias Opialla, Dr. Matthias Pietzke, Nadine Royla, Dr. Christin Zasada – thank you so much. Especially I am thankful to Dr. Jennifer Kirwan and Dr. Tobias Opialla for their tremendous help with this thesis: dear Jennifer and Tobi, I would be completely lost without your help, thank you again and again and again.

Of course, I am grateful for the support from my grad school – Berlin School of Integrative Oncology – and specifically to Prof. Dr. Clemens Schmitt and Dr. Eleanor Eife-Horn.

I have to also acknowledge my friends: Sasha Bajev and Natasha Bajeva, Dr. Dima Boiko, Nastya Boltengagen, Ilya Borodkin, Vasya Desyatov and Natasha Desyatova, Merve Erdem, Andrei Filipchyk, Maggy Herzog, Kshu Galilova, Dr. Anton Kondakov, Dr. Benedikt Kortüm, Dr. Akihiro Maeda, Polya Peeva, Dr. Dima Svistunenko, Dr. Dasha Federova and Uta Wegner, who always believed in me and supported with words and actions. I love you, guys.

So many thanks goes to my better half – my twin sister Anya Vv, for always standing by my side, backing me up when needed and always believe in in me. However, of course, the whole Vv Family was always there for me: Dasha, Mom, Aunt, and Grandpa.

I would like to close the Acknowledgments part by thanking the person, without whom my life, both scientific and non-scientific, would be completely different. Dear Andy (Dr. Andrew Anthony Amoscato, to be precise), thank you so much for being the best teacher and friend one can imagine. One day long time ago you told, that I could consider myself the “third Amoscato twin”, and if has been so ever since. The great teacher inspires, and that is exactly what you did: you basically introduced me to the world of science and mass spectrometry, inspired me to stay in science, to go for Ph.D. hunt, and now to keep going further, and I know that you will always be there, even when I am to swim with sharks. When I count my blessings, I count you twice.

Table of Contents

Abbreviations

2D-PAGE	Two dimensional polyacrylamide gel electrophoresis
ACLY, Acly	ATP-citrate synthase
ACO1, Aco1	Cytoplasmic aconitate hydratase
ACO2, Aco2	Aconitate hydratase, mitochondrial
ACSL1, Acsl1	Long-chain-fatty-acid--CoA ligase 1
ACSL3, Acsl3	Long-chain-fatty-acid--CoA ligase 3
ACSL4, Acsl4	Long-chain-fatty-acid--CoA ligase 4
ACSL5, Acsl5	Long-chain-fatty-acid--CoA ligase 5
ACSL6, Acsl6	Long-chain-fatty-acid--CoA ligase 6
ADP	Adenosine diphosphate
AFP	Alpha-fetoprotein
AFP-L3	Fucosylated alpha-fetoprotein
ALDOA, Aldoa	Fructose-bisphosphate aldolase A
ALDOB, Aldob	Fructose-bisphosphate aldolase B
ALDOC, Aldoc	Fructose-bisphosphate aldolase C
ALT	Alanine transaminase
AMP	Adenosine monophosphate
AST	Aspartate transaminase
ATP	Adenosine triphosphate
B2M	Beta-2-microglobulin
BMI	Body Mass Index
CCM	Central carbon metabolism
CoA	Coenzyme A
CS, Cs	Citrate synthase, mitochondrial
CTNNB1	Catenin Beta 1 (β-catenin)
CTNNB1	Computerized tomography
DCP	Des-gamma carboxyprothrombin
DLAT, Dlat	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial
DLD, Dld	Dihydrolipoyl dehydrogenase, mitochondrial
DLST, Dlst	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EB1	APC-binding protein EB1
ENO1, Eno1	Alpha-enolase
ENO2, Eno2	Gamma-enolase
ENO3, Eno3	Beta-enolase
FA	Fatty acid
FASN, Fasn	Fatty acid synthase
FBP1, Fbp1	Fructose-1,6-bisphosphatase 1
FBP2, Fbp2	Fructose-1,6-bisphosphatase isozyme 2

FDR	False discovery rate
FH, Fh	Fumarate hydratase, mitochondrial
G6PC, G6pc	Glucose-6-phosphatase
G6PD, G6pdx	Glucose-6-phosphate 1-dehydrogenase
GAPDH, Gapdh	Glyceraldehyde-3-phosphate dehydrogenase
GC	Gas chromatography
GCK, Gck	Glucokinase
GGT	Gamma-glutamyltransferase
GLS, Gls	Glutaminase kidney isoform, mitochondrial
GLS2, Gls2	Glutaminase liver isoform, mitochondrial
GLUD1, Glud1	Glutamate dehydrogenase 1, mitochondrial
GLUD2	Glutamate dehydrogenase 2, mitochondrial
GO	Gene ontology
GPI, Gpi	Glucose-6-phosphate isomerase
GPT, Gpt	Alanine aminotransferase 1
GPT2, Gpt2	Alanine aminotransferase 2
GYS1, Gys1	Glycogen [starch] synthase, muscle
GYS2, Gys2	Glycogen [starch] synthase, liver
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HK1, Hk1	Hexokinase-1
HK2, Hk2	Hexokinase-2
HK3, Hk3	Hexokinase-3
HP	Haptoglobin
HSP	Heat Shock Protein
IAA	Iodoacetamide
IDH1, Idh1	Isocitrate dehydrogenase [NADP] cytoplasmic
IDH2, Idh2	Isocitrate dehydrogenase [NADP], mitochondrial
IDH3A, Idh3a	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial
IDH3B, Idh3b	Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial
IDH3G, Idh3g	Isocitrate dehydrogenase [NAD] subunit gamma, mitochondrial
IGF2	Insulin-like growth factor 2
IGHG1	Ig gamma-1 chain C region
LC	Liquid Chromatography
LDHA, Ldha	L-lactate dehydrogenase A chain
LDHB, Ldhb	L-lactate dehydrogenase B chain
LDHC, Ldhc	L-lactate dehydrogenase C chain
LDHD, Ldhd	Probable D-lactate dehydrogenase, mitochondrial
LFQ	Lable-free quantification
m/z	mass/charge
MCW	Methanol-chloroform-water
MDH1, Mdh1	Malate dehydrogenase, cytoplasmic

MDH2, Mdh2	Malate dehydrogenase, mitochondrial
ME1, Me1	NADP-dependent malic enzyme
ME2, Me2	NAD-dependent malic enzyme, mitochondrial
ME3, Me3	NADP-dependent malic enzyme, mitochondrial
MPC1, Mpc1	Mitochondrial pyruvate carrier 1
MPC2, Mpc2	Mitochondrial pyruvate carrier 2
MS	Mass spectrometry
MSTFA	N-methyl-N-[trimethylsilyl] trifluoroacetamide
NAD+	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide phosphate
NASH	Non-alcoholic steatohep
NMR	Nuclear magnetic resonance
OGDH	Oxoglutarate dehydrogenase
OGDH, Ogdh	2-oxoglutarate dehydrogenase, mitochondrial
OGDHL, Ogdhl	2-oxoglutarate dehydrogenase-like, mitochondrial
PC, Pc	Pyruvate carboxylase, mitochondrial
PCK1, Pck1	Phosphoenolpyruvate carboxykinase, cytosolic [GTP]
PCK2, Pck2	Phosphoenolpyruvate carboxykinase [GTP], mitochondrial
PDHA1, Pdha1	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial
PDHA2, Pdha2	Pyruvate dehydrogenase E1 component subunit alpha, testis-specific form, mitochondrial
PDHB, Pdhb	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial
PDHX, Pdhx	Pyruvate dehydrogenase protein X component, mitochondrial
PFKFB1, Pfkfb1	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 1; 6-phosphofructo-2-kinase;Fructose-2,6-bisphosphatase
PFKFB2, Pfkfb2	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 2; 6-phosphofructo-2-kinase;Fructose-2,6-bisphosphatase
PFKL, Pfkl	ATP-dependent 6-phosphofructokinase, liver type
PFKM, Pfkm	ATP-dependent 6-phosphofructokinase, muscle type
PFKP, Pfkp	ATP-dependent 6-phosphofructokinase, platelet type
PGAM1, Pgam1	Phosphoglycerate mutase 1
PGAM2, Pgam2	Phosphoglycerate mutase 2
PGAM5, Pgam5	Serine/threonine-protein phosphatase PGAM5, mitochondrial
PGD, Pgd	6-phosphogluconate dehydrogenase, decarboxylating
PGK1, Pgk1	Phosphoglycerate kinase 1
PGK2, Pgk2	Phosphoglycerate kinase 2
PGM	Personal Genome Machine
PGM1, Pgm1	Phosphoglucomutase-1
PGM2, Pgm2	Phosphoglucomutase-2
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha
PIVKA-II	Protein Induced by Vitamin K Absence/antagonist-II
PKLR, Pklr	Pyruvate kinase PKLR
PKM, Pkm	Pyruvate kinase PKM

PON1	Serum paraoxonase/arylesterase 1
PPP	Pentose phosphate pathway
PYGL, Pygl	Glycogen phosphorylase, liver form
SDHA, Sdha	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial
SDHB, Sdhb	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial
SDHC, Sdhc	Succinate dehydrogenase cytochrome b560 subunit, mitochondrial
SDHD, Sdhd	Succinate dehydrogenase [ubiquinone] cytochrome b small subunit, mitochondrial
SNP	Single-nucleotide polymorphism
SOX9	Sex determining region Y boxes-9
SUCLA2, Sucla2	Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial
SUCLG1, Suclg1	Succinyl-CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial
SUCLG2, Suclg2	Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial
TALDO1, Taldo1	Transaldolase
TCA	Tricarboxylic acid
TCF1	Hepatocyte nuclear factor 1
TFA	Trifluoroacetic acid
TKT, Tkt	Transketolase
TKTL1, Tktl1	Transketolase-like protein 1
TKTL2, Tktl2	Transketolase-like protein 2
TOF	Time-of-flight
TP53	Tumor protein tp53
TPI1, Tpi1	Triosephosphate isomerase
UHPLC	Ultra-High Performance Liquid Chromatography

Introduction

In the introduction I present general information about hepatocellular carcinoma (HCC), the details about prognosis, epidemiology, risk factors, and classification. Further, I discuss the clinical picture, diagnostic HCC markers and treatment possibilities. In the Introduction, molecular mechanisms of central carbon metabolism are described, including its liver-specific features and regulation. Here I also introduce the HCC-specific changes of CCM and systems biology and translational medicine approaches to this problem. The Introduction finishes with the aims of the present work.

Hepatocellular carcinoma

Et Tumor, Brute?

-Bethesda Softworks, *Fallout 4, 2015*

HCC is a malignant liver tumor with hepatocellular differentiation (Fritz, et al., 2000). Diverse histological subtypes of HCC are described (Bosman, Carneiro, Hruban, & Theise, 2010; Deepali, 2015; Fritz, et al., 2000), however, all of them express common trends. Most common patterns are trabecular with 4 or more cells surrounded by a layer of flattened endothelial cells. Sinusoidal vessels surrounding tumor cells are an important feature of HCC, as well as scanty stroma, and vascular invasion (a detailed histological picture can be found in l Appendix 1) (Deepali, 2015).

Prognosis

The prognosis for patients with HCC is generally very poor. Statistical studies report median of less than 5-years survival of HCC patients after the diagnosis (Altekruse, McGlynn, & Reichman, 2009; Bosman, et al., 2010; Fritz, et al., 2000; Greten, et al., 2005; Lee, et al., 2006). This value may vary depending on the etiology or treatment, for example median survival for HBV[1]-related HCC in less than 16 months (Nguyen, Law, & Dore, 2009).

Epidemiology

HCC is the most common type of liver cancer, accounting for 90% of all liver cancers (Jelic & Sotiropoulos, 2010). It is a major public health problem, accounting for approximately 745 000 deaths worldwide in 2012 according to the data from the International Agency for Research on

[1] HBV - hepatitis B virus

Cancer (IARC, 2012). The number of diagnosed cases is similar to the number of deaths attributed to HCC worldwide, which highlights that diagnostic tools and treatment methods are currently limited, since proper diagnosis is often made in a disease state where treatment options focus on palliative care.

Risk factors

Several risk factors are described by researchers and doctors. In most cases, HCC is associated with chronic liver injury of different etiology (HBV, HCV[1] and alcoholic cirrhosis). Chronic infection with HBV in the setting of cirrhosis increases the risk of HCC from 5- to 100-fold (El-Serag, 2012). Statistically speaking, 5–30% of individuals with HCV infection develop chronic liver disease, 30% of them progress to cirrhosis. 1–2% of these patients per year without proper treatment and elimination of damaging factors will develop HCC. Median age at diagnosis is between 50 and 60 years (Jelic & Sotiropoulos, 2010). Overall, patients with cirrhosis are at highest risk of developing HCC (Forner, Llovet, & Bruix, 2012). HCC is four to eight times more common in men (Jelic & Sotiropoulos, 2010). Chiesa et al. in 2000 discovered that 87% of the cases of HCC occurring in Brescia, northern Italy, developed in a cirrhotic liver, including most of those attributable to alcohol intake (Chiesa, et al., 2000). Ukawa et al. discovered that 26% of 1245 HCC patients from BioBank Japan suffered from alcoholism (Ukawa, et al., 2017).

Additionally, other toxins can increase the risk of developing HCC, for example aflatoxin. Aflatoxin, a contaminant produced by the fungi *Aspergillus flavus* and *Aspergillus parasiticus* in maize and nuts, is a known human liver carcinogen (Liu & Wu, 2010). This factor is a good example for territory-dependent HCC risk factors. Most cases are detected in sub-Saharan Africa, Southeast Asia and China, where populations suffer from both high HBV and food aflatoxin exposure (Liu & Wu, 2010).

The above mentioned risk factors have been proven to strongly correlate to the HCC incidence. Human liver is a primary organ of detoxification. Liver enzymes catalyze the oxidation, hydrolysis and reduction (phase I) of chemical molecules and drugs, and their conjugation (phase II) (Grant, 1991). However, several chemicals could be converted to more toxic substances and metabolites, causing toxic effects on liver and whole organism. The liver is extremely sensitive to chemical- and viral-induced carcinogenesis.

[1] HCV – hepatitis C virus

Steatosis and non-alcoholic steatohepatitis (NASH) - early stages of non-alcoholic fatty liver disease (NAFLD) - are associated with abdominal obesity and type 2 diabetes (Margini & Dufour, 2016). For example, 95-100% of patients with BMI[1] over 30 kg/m2 are diagnosed with steatosis, and 20-47% with NASH (Ludwig, Viggiano, McGill, & Oh, 1980; Powell, et al., 1990). Patients with type 2 diabetes or with defective glucose tolerance are diagnosed with steatosis in 60% of cases, and with NASH in 15% of cases. Patients with lipid metabolism defects are diagnosed with steatosis and NASH in 20-80% of cases (Roytberg & Strutynskiy, 2013). The epidemic of obesity has been correlated to an increased risk of various types of cancer (Vucenik & Stains, 2012). According to the study by Ukawa et al. on the BioBank Japan project 22% of 1245 HCC patients were obese (BMI ≥30 kg/m^2) (Ukawa, et al., 2017).

Gene polymorphisms and gene-environment interactions may also be risk factors for HCC, particularly genes that code for metabolic enzymes (glutathione-S-transferase, epoxide hydrolase, cytochrome p4502E1) or DNA repair enzymes (XRCC1, UDP-glucuronosyltransferase 1A7) (Borentain, et al., 2007; Degli Esposti, et al., 2012).

Due to the variety of etiology described above, many HCC-associated molecular alterations have been detected and no universal molecular signature is definitively associated with all hepatic tumors, neither in humans nor in experimental animals (Degli Esposti, et al., 2012; Pei, Zhang, Renault, & Zhang, 2009).

It is rather difficult to distinguish the risk factors of HCC from its cause and from the associated diseases. Additionally, multiple risk factors (or causes) can be present in one HCC case. Therefore, the classification based on the cause of HCC cannot be considered reliable due to its inability to provide unambiguous grouping.

Classification

Mutation-based classification, suggested by Boyault et al. (Boyault, et al., 2007), does not rely exclusively on risk factors/causes. It is mostly based on involved biological pathways, gene mutations and methylation, while clinical information does not play a major role or sometimes is not considered at all. Six major groups of different types of HCC and its genetic and clinical (when available) alterations are described (Figure 1).

The first (G1) group of tumors is characterized by a low copy number of HBV and by an overexpression of the following genes: sex determining region Y boxes-9 (SOX9), insulin-like

[1] BMI – Body Mass Index

growth factor 2 (IGF2), alpha-fetoprotein (AFP) and other genes which are expressed in fetal liver. G2 includes HCCs with a high copy number of HBV and mutations in phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PIK3CA) and tumor protein tp53 (TP53). In G1 and G2 groups AKT/PKB signaling pathway activation was detected. G3 contains tumors with mutated genes TP53 and an overexpression of genes responsible for the cell cycle control. In particular, the DNA replication licensing factors MCM2, 3 and 6 are overexpressed**.** G4 is described as a heterogeneous tumor group including hepatocyte nuclear factor 1 (TCF1)-mutated hepatocellular adenomas and carcinomas. G5's distinguishing feature is the presence of catenin mutations that lead to Wnt pathway activation. The beta-catenin target genes in the liver are overexpressed. The picture for the G6 group is similar to G5 but additionally includes tumors with satellite nodules, higher activation of the Wnt pathway, overexpression of beta-catenin target genes and E-cadherin underexpression. Boyault et al. came to the conclusion that approximately 50% of the tumors were related to the activation of either AKT or WNT pathways.

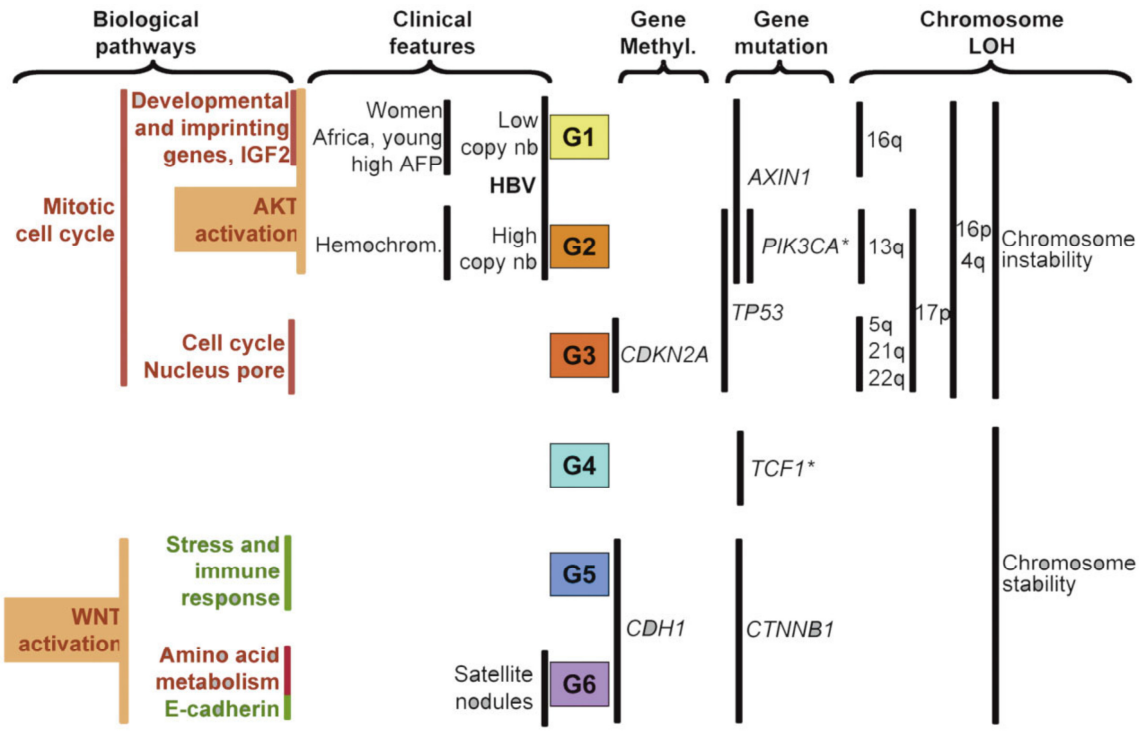


Figure 1. G1-G6 HCC groups. Schematic of the different HCC subgroups defined by transcriptome analysis with their related clinical and genetic pathways. G1 to G6 are the subgroups of HCCs defined by transcriptome analysis. Vertical lines indicate significantly associated features. Red and green primarily indicate over- and underexpressed genes, respectively, in that particular functional

category. LOH, loss of heterozygosity; Hemochrom., hemochromatosis; AFP, alpha-fetoprotein; HBV, hepatitis B virus; * rare feature (Boyault, et al., 2007).

This classification covers most of the known genetic alterations, present in HCC and includes a few clinical features detected in patients. It is interesting that G3-G6 groups do not show any specific clinical component. In the present work, I used the above mentioned classification as a starting point for the genetic analysis. Currently, genomic alterations are mostly used for research purposes, but are not widely applied in clinics for diagnosis or treatment of HCC.

Diagnosis

Clinical picture

The strategy of HCC diagnostics follows an established protocol. However, it can vary in different cases. Usually patients with chronic liver damage have several symptoms, such as jaundice, tiredness (fatigue), itch, discomfort over the liver, and red or pink blotchiness of the hands. Considering these symptoms doctors suggest several diagnostics tools: ultrasound, measurement of liver enzymes (the most common specific elevated liver enzymes are alanine transaminase (ALT), aspartate transaminase (AST), and gamma-glutamyltransferase (GGT), and they are currently considered to be non-specific for HCC), computerized tomography (CT) and, to confirm the suggested diagnosis, histological examination of a biopsy sample. As a general rule, patients with either hepatitis B or C or cirrhosis are recommended for regular ultrasound checkups (Roytberg & Strutynskiy, 2013).

HCC markers

Testing for specific HCC markers is considered a rather questionable method at the moment due to its lack of specificity. For a long time alpha-fetoprotein (AFP) was used as a common marker for HCC. Discovered about half a century ago AFP is the major serum fetal protein in mammals (Abelev, Perova, Khramkova, Postnikova, & Irlin, 1963; Bergstrand & Czar, 1956). It is mostly produced during the lifetime of the fetus by liver hepatocytes. If the hepatocyte undergoes cancerous changes it starts to produce more AFP. However, AFP has been proven by several authors to be non-specific as an HCC marker. According to research by Paul et al. AFP shows no significant elevation in patients with cirrhosis compared to patients with cirrhosis and HCC (Paul, et al., 2007). Gupta and his colleagues came to the same conclusion for patients with hepatitis C (Gupta, Bent, & Kohlwes, 2003). Tomomi Masuda and Eiji Miyoshi noted, that levels of AFP, GP73,

HSP70, and glypican 3 can also correspond to chronic inflammation and stress (Masuda & Miyoshi, 2011).

Several authors have claimed that fucosylated AFP (AFP-L3) has better diagnostic value (Luk & Liu, 2011; Masuda & Miyoshi, 2011; Zinkin, et al., 2008).

The level of Protein Induced by Vitamin K Absence/antagonist-II (PIVKA-II) (also known as des-gamma carboxyprothrombin (DCP)) can change in the presence of HCC (Bertino, et al., 2013; Masuda & Miyoshi, 2011; Zinkin, et al., 2008).

Heat Shock Protein 70 (HSP70) and HSP27 are of special relevance in human cancer because they inhibit apoptosis. HSP70 could be a sensitive marker for the differential diagnosis of early HCC from precancerous lesions or non-cancerous liver, a difficult distinction for pathologists due to the very well-differentiated histology with little atypia in early HCC (Masuda & Miyoshi, 2011).

Those and many other potential markers can be found in literature, but none of the suggested markers is used nowadays as a 'Gold Standard' for the differential HCC diagnosis. Thus, current biomarker-based diagnostics of HCC requires a lot of improvement. One of the aims of this thesis is to obtain more specific knowledge about possible novel HCC markers.

Treatment

Main treatment options, according to (Roytberg & Strutynskiy, 2013), include:

- Liver transplantation (Chang, et al., 2017);
- Partial surgical resection to remove a tumor and surrounding liver tissue, retaining normal liver tissue (J. Y. Shen, et al., 2016);
- Interventional radiology (various embolizations (Andreana, et al., 2012; Lobo, et al., 2016), radiofrequency ablation (damages tumor with local heating) (Clasen, et al., 2014; Ueno, et al., 2015), selective internal radiation therapy (Hilgard, et al., 2009; Lambert & Van De Wiele, 2009), percutaneous ethanol injection (Huang, et al., 2015; Yang, et al., 2015) etc.);
- Sorafenib (receptor tyrosine kinase inhibitor) is occasionally used in patients with advanced HCC (Keating, 2017; Llovet, et al., 2008; A. Shen, et al., 2013). Sorafenib

inhibits tumor-cell proliferation and tumor angiogenesis, and increases the rate of apoptosis in other tumor models.

Despite obvious diversity of causes, background conditions, genetic alterations, clinical pictures, and histological types of HCC, current treatment options are non-selective in the majority of cases. Possibly, this non-specific approach leads to an overall poor prognosis of HCC. Deep analysis of various metabolic phenotypes of different HCC cases could potentially shed the light on discrepant classification and treatment options of this severe disease.

Molecular mechanisms

Healthy liver tissue plays a major role in carbohydrate metabolism: it synthesizes and stores glycogen, releases glucose into the blood by glycogenolysis, performs glucose *de novo* synthesis. Some of these processes are part of central carbon metabolism (CCM). CCM is a biochemical network of enzymes and metabolites of cells, representing the most basic aspect of life (Kanehisa, Furumichi, Tanabe, Sato, & Morishima, 2017; Kanehisa & Goto, 2000; Kanehisa, Sato, Kawashima, Furumichi, & Tanabe, 2016; Sudarsan, Dethlefsen, Blank, Siemann-Herzberg, & Schmid, 2014). The classic textbook knowledge of CCM includes the Embden-Meyerhof-Parnas pathway (glycolysis), the TCA cycle, and the pentose phosphate pathway (Nelson, Nelson, Lehninger, & Cox, 2008). Additional pathways connected to CCM, such as *do novo* synthesis of glucose and glycogen, and glutaminolysis, are analyzed in this work as they participate in crucial liver functions. The scheme of these processes is presented on Figure 2 (full names of enzymes and metabolites for both human and mouse species see in Appendix 2). It includes major enzymes participating in CCM and its isoforms. These different forms of proteins may be produced from different genes, or from the same gene by alternative splicing (MeSH, 1999), and have similar functions. Structurally related forms of an enzyme are called isoenzymes. Each isoenzyme has the same mechanism and classification, but differs in its chemical, physical, or immunological characteristics (MeSH, 1999). In normal conditions one isoform or isoenzyme is quantitatively prevalent over another, and their ratio can be changed in various biological processes, such as cancer development. These processes is called an isoform switch (due to historical reasons). Understanding the interactions between CCM members in various tissues under healthy and pathological conditions is of primary interest in systems biology.

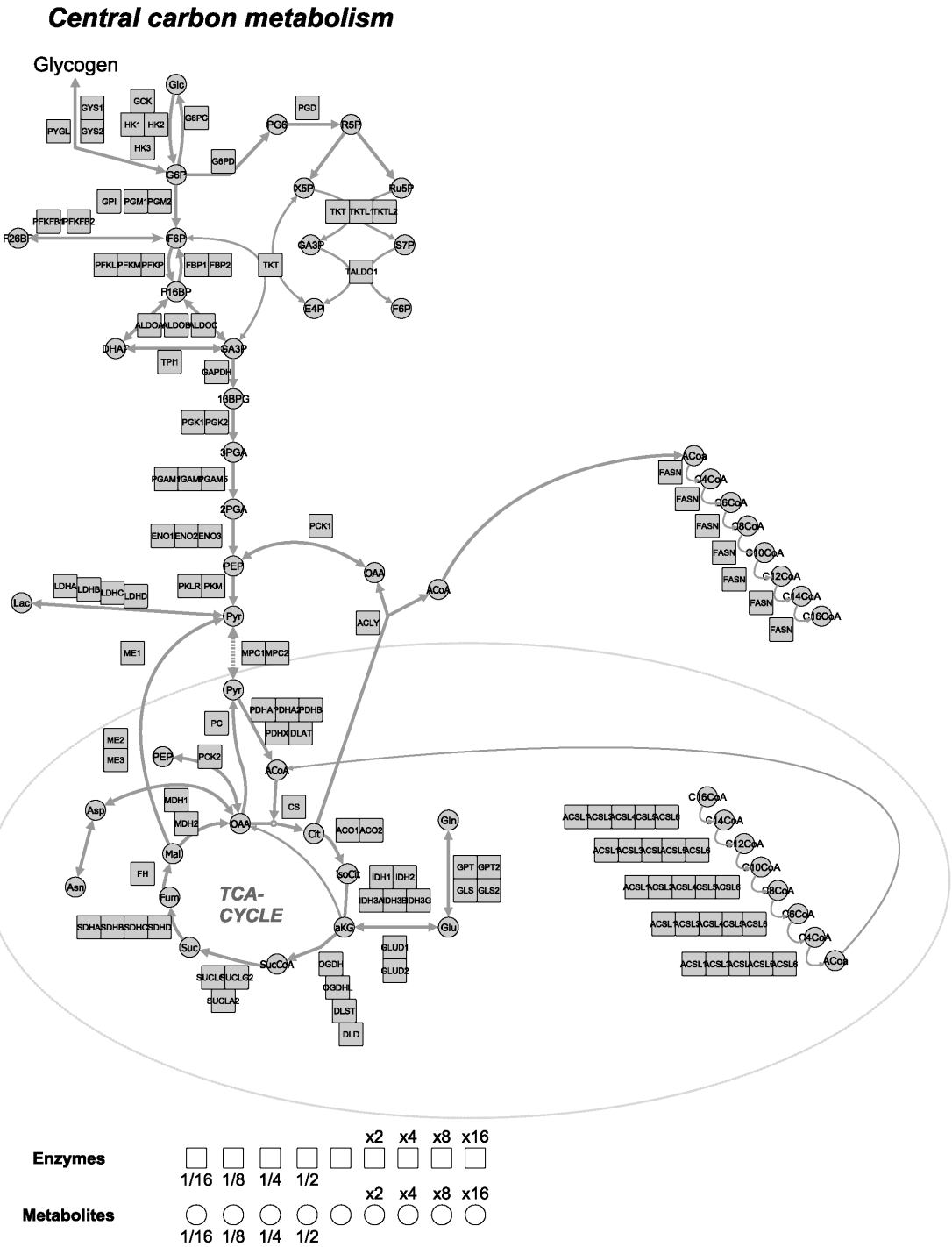


Figure 2. Scheme of central carbon metabolism (CCM). Visualization of major CCM enzymes and its isoforms and subunits, and metabolites. Enzymes are visualized with squares, metabolites with circles, mitochondrion membrane is indicated with the grey line. For the full names of enzymes and metabolites for both human and mouse species see in Appendix 2.

Regulation of CCM

As some of central carbon metabolism (CCM) processes run in opposite directions, they have to be strictly regulated. If glycolysis and gluconeogenesis were active both at the same time, glucose would be converted to pyruvate, and then back to glucose, which would lead to futile ATP[1] consumption (Boiteux & Hess, 1981). While this cycle exists in nature to produce heat in the brown adipose tissue, in terms of liver metabolism this would be a drastic waste of energy. Therefore, in liver tight regulation is required to facilitate liver adjustment to constantly changing circumstances.

Glycolysis in liver is mainly regulated by four major enzymes: GCK, HK, PFK, and PK. GCK has a special role in liver: releasing glucose to the blood when blood glucose is low, and taking up and metabolizing glucose when the blood glucose level is high. PFK is formed by different combinations of three types of subunits: platelet (P), muscle (M), and liver and red blood cells (LR). PFK is allosterically inhibited by ATP and citrate, and activated by fructose-2,6-bisphosphate. PK is inhibited by ATP and acetyl-CoA[2], and PKLR is inhibited by c-AMP[3]. On the other hand fructose-1,6-bisphosphate activates PKM. Gluconeogenesis is regulated at the level of PCK1, which is activated by acetyl-CoA, and by FBP1, which is inhibited by fructose-2,6-bisphosphate and AMP. Glucagon raises cAMP level in the cell, and inhibits fructose-2,6-bisphosphate. This leads to the inhibition of glycolysis and stimulation of gluconeogenesis. Insulin on its turn increases the level of fructose-2,6-bisphosphate, which leads to an opposite effect: stimulation of glycolysis and inhibition of gluconeogenesis (Elliott & Elliott, 2005; Nelson, et al., 2008).

The TCA cycle is controlled by the conversion ratio of pyruvate to acetyl-CoA, and by the flux through CS, OGDH[4], and IDH. These fluxes are regulated by the concentrations of substrates and products: NADH[5] and ATP are inhibitors, and NAD+[6] and ADP[7] are activators. The key enzyme PDH

[1] ATP - Adenosine triphosphate

[2] CoA – Coenzyme A

[3] AMP - Adenosine monophosphate

[4] OGDH – Oxoglutarate dehydrogenase

[5] NADH - Nicotinamide adenine dinucleotide phosphate

[6] NAD+ - Nicotinamide adenine dinucleotide

[7] ADP - Adenosine diphosphate

complex is allosterically inhibited by several metabolites including acetyl-CoA, NADH, ATP and fatty acids, while the PDH complex is activated by AMP, NAD+ and CoA (Elliott & Elliott, 2005; Nelson, et al., 2008).

Changes of CCM in cancer

The above described liver metabolism and its regulation is valid for the healthy liver. However, the impairment of metabolism during liver cancer has not yet been investigated in a global way. The vast majority of publications are devoted to metabolism of other cancer types. Under normal conditions most healthy cells have a low rate of glycolysis which is followed by further oxidation of pyruvate in the mitochondria. According to Otto Warburg's findings in 1956 (Warburg, 1956), a hallmark of cancer cells is anaerobic glycolysis even under aerobic conditions. This behavior, involving increased glucose consumption with consequent excessive production of lactate, is called the Warburg effect. Further studies of the Warburg effect in various cancer types have addressed the efficiency of cancer cell metabolism. Even though mitochondrion metabolism does not work at full capacity, the metabolites produced from nutrients in CCM can be taken up and incorporated into the nucleotides, amino acids, and lipids (Vander Heiden, Cantley, & Thompson, 2009). However, liver metabolism is different from the metabolism of other tissues and organs, it has various metabolic functions, and regulation of these functions is liver-specific. One of the main aims of this work is to find out whether the samples analyzed in the present work show a Warburg effect-like behavior, and evaluate the role of major regulator enzymes of CCM and others enzymes participating in liver metabolism.

The diversity of metabolic phenotypes and metabolic reprogramming of HCC has not yet been studied with a systematic approach. Several authors, however, addressed this subject in a focused way. For example, Yu and colleagues in 2015 suggest kidney-type glutaminase (GLS1) as a biomarker for pathologic diagnosis and prognosis of HCC (Yu, et al., 2015). The authors studied an altered glutaminolysis in HCC, but did not perform a full CCM analysis. A conceptually similar approach was applied by Hirata et al. in their work devoted to fructose-1,6-bisphosphatase level changes during tumor progression of HCC (Hirata, et al., 2016). A more global and ambitious attempt at analysis of the tissue metabolome of HCC was published by Beyoğlu and colleagues in 2013. These authors studied tumor energy metabolism and the role of the above mentioned transcriptomic classification (Beyoglu, et al., 2013). While the authors indeed found some differences in various types of HCC (based on G1-G6 classification), their cohort and human tissue material choices had limitations, which will be properly discussed in this thesis later. While recognizing the scientific significance of those very thorough and detailed works, it is important to

point out a lack of global systematic studies of the whole CCM as a process in the field of HCC research. The deficit of fundamental knowledge of processes taking place during HCC development may lead to a lack of diagnostic and treatment options, and therefore to a poor disease prognosis. Elucidation of HCC-specific key features of metabolism will lead to a better understanding of this disease, and potentially point to additional diagnostic and treatment options.

Systems biology approach

> Systems biology...is about putting together rather than taking apart, integration rather than reduction…
>
> -Denis Noble, *The music of life: Biology beyond the genome, 2006*

To evaluate the metabolic phenotype of HCC a systematic approach is required (R. Z. Shang, Qu, & Wang, 2016), and systems biology is ideally suited to work on this task (Tian, Price, & Hood, 2012; Willekens, et al., 2009). Systems biology studies the interactions between various components of biological systems and how these interactions affect the behavior of that system (Snoep & Westerhoff, 2005). For example, proteome and metabolome studies uncovers the *in situ* phenotypes of various tissues (Yizhak, Benyamini, Liebermeister, Ruppin, & Shlomi, 2010). Prior to mass spectrometry, proteomics approaches mostly included western blotting (Burnette, 1981; Paoletti & Chang, 2000) and enzymatic essays (Reisch & Elpeleg, 2007; Tietze, 1969). While these methods are reliable and specific, they are rather time-consuming, which reduces the possibility of comparatively fast thorough analysis of large amount of proteins. Two dimensional polyacrylamide gel electrophoresis (2D-PAGE) allows the analysis of hundreds to thousands of proteins, but its pipeline does not include automated high throughput analysis (Berth, Moser, Kolbe, & Bernhardt, 2007; Klose, 1975; O'Farrell, 1975). By contrast, modern shotgun mass spectrometry proteome analysis grants the possibility to reliably detect and identify thousands of proteins in tissue sample as small as 1 mg (Guo, et al., 2015; Jacobs, et al., 2005).

Metabolomic methods allow users to analyze various metabolites in living systems. Metabolomics studies are performed mostly with several methods, such as nuclear magnetic resonance (NMR) (Cohen, Shulman, & McLaughlin, 1979; Lin, Wu, Tjeerdema, & Viant, 2007; Martínez-Granados, et al., 2006), various assays (Gough, Armour, & Baker, 1997; Tolosa, Malak,

Raob, & Lakowicz, 1997)**,** and analytical instruments (such as Seahorse XF technology by Agilent (Pandya, et al., 2016)), and MS. NMR is reproducible, however the sensitivity is an issue. Seahorse XF technology allows to measure only few products of the intracellular reactions. The spectrum of metabolites analyzed by these methods and assays is very limited compared to the possibilities of MS. In-house MS associated bioinformatics pipelines developed in the Kempa group allow us not only to detect and identify, but also to potentially quantify more than a hundred metabolites in a single sample (Kuich, Hoffmann, & Kempa, 2014).

Genome analysis can be performed in using both non-targeted (whole genome sequencing (Fujimoto, et al., 2012; Ng & Kirkness, 2010)) and targeted methods. The latter is achieved by panel sequencing (Easton, et al., 2015; Kammermeier, et al., 2014), where the researched focuses on analysis of known mutations. It allows the detection of mutations (for example, main somatic cancer-associated mutations) in the genome and avoids massive data sets, in contrary to whole genome sequencing. In the present work, I used panel sequencing to identify the main cancer-related somatic mutations in human liver tissue. Combined with proteomics and metabolomics, analysis of mutations could allow the evaluation of links between mutations and metabolic phenotypes.

Thus, I focused on proteomic and metabolomic analyses as well as genotyping for a better understanding of the mechanisms which cause changes in the liver during progression from a normal state to HCC.

Aims

Current work aims to answer the following questions:

- **What is the role of metabolism in HCC development?**
- **Are the metabolic changes in HCC-affected liver tissue mirrored in serum samples?**
- **Which differences can be detected in the metabolomic, proteomic, and genomic profiles of liver tissues with developed HCC compared to non-cancerous liver tissue?**
- **How heterogeneous are the CCM profiles of various HCC samples across multiple patients?**
- **How are the proteomic and metabolomic profiles and perturbations in HCC interconnected and how does this relate to the genetic mutations status of HCC tissue?**
- **Is there a way to predict the patient's prognosis?**

Mouse model

The description of the model we used in our research was published by Nathalie Dubois and her colleagues in 1991 (Dubois, et al., 1991). The authors developed a transgenic ASV-B mouse HCC model with the presence of lung metastasis in 10% of the cases. The ASV-B mouse model was established in the B6D2F1 mouse, which is a cross between a C57BL/6J female and a DBA/2J male. A precise targeting of the SV40 T early region expression in the liver of transgenic mice was obtained using 700 bp of the antithrombin III regulatory sequence to control oncogene expression. In the strain expressing the highest level of SV40 large T antigen (Tag), the incidence of hepatocarcinoma was 100% (Dubois, et al., 1991). The progress of HCC was reproducible and characterized by the development of cytolysis by the 4th week, when no morphological and histological modifications were visible. Hepatic cytolysis can be indicated by measuring the ornnithine transcarbamylase (OTC) activity in the serum of mice (Dubois, et al., 1991). The OTC level increases drastically in plasma after hepatic cell lysis (Feigelson, Pecau, Cathelineau, & Navarro, 1975). A pre-neoplastic state was marked by a progression from hyperplasia to proliferative nodules composed of highly differentiated cells exhibiting a high Tag expression, which could elicit tumor formation in nude mice and could proliferate *in vitro*. In 10% of the cases, hepatocellular carcinoma was associated with lung metastasis (Dubois, et al., 1991).

The ASV-B mouse model used shows dysplastic nodule formation at the age of 8 weeks. By the 12th week, adenomas in the liver can be observed and differentiated HCC develops in these mice by the 16th week. The cohort for the current study was formed from 7 mice at 16 weeks old with fully developed HCC and 7 C57BL/6J mice of the same age as healthy controls. The C57BL/6J mice had to be used as healthy control, because all ASV-B mice (both males and females) are genetically modified thus cannot be considered healthy.

Materials and methods

Mice

Mouse model

The ASV-B model described by (Dubois, et al., 1991) was used in this study. The cohort included 7 ASV-B mice at 16 weeks old and 7 healthy C57BL/6J at 16 weeks old as a healthy control. Both male and female mice from ASV-B model have modified genetic background, thus the healthy control had to be obtained from a different mouse strain.

All mice were maintained under standard conditions at the animal facilities in Berlin (Charité) and Aachen (Institut für Versuchstierkunde, University Hospital Aachen). Animal procedures were performed in accordance to approved protocols (Landesamt für Gesundheit und Soziales, Berlin (0024/12) and Landesamt für Natur, Umwelt und Verbraucherschutz, Düsseldorf (84-02.04.2015.A344, AZ84-02.04.2016.A018 and 84-02.04.2015.A216)) and followed recommendations for proper care and use of laboratory animals.

Material overview

The liver tissue obtained from the mice was used for proteomic and metabolomic analysis. A full description of the method development is described in Results section. The amount of material used for proteomics was: 4.4-9.1 mg, for metabolomics: 3.9-8.9 mg; and for test extraction of DNA: 4.9-8.2 mg. DNA was extracted and quantified in order to estimate the amount of material which could be obtained by this method. However, in this study we have not performed the genomics analysis of the mouse model, because all mice have the same known genetic background.

Proteomics

The samples were lysed in urea buffer using a Tissue Lyser (Precellys 24 lysis and homogenization, Bertin Technologies, France) and zirconium beads (5000 rpm, 20 second cycle time, 2 cycles).

In-solution digestion

Proteins were reduced in 2mM dithiothreitol (DTT) for 30 minutes at 25 °C and free cysteines were alkylated (to prevent the formation of disulfide bridges) in 11 mM iodoacetamide (IAA) for 20 minutes at room temperature in darkness. LysC digestion was performed by adding LysC (Wako) in a ratio 1:40 (w/w) to the sample and incubating it for 18 hours under gentle

shaking at 30°C. After LysC digestion, the samples were diluted 4-fold with 50 mM ammonium bicarbonate solution, 7 µL of immobilized trypsin (Applied Biosystems) was added and samples were incubated for 4 hours on a rotator at 30 °C (Rotator SB3, Stuart, UK). Digestion was stopped by acidification with 10 µL of trifluoroacetic acid (TFA) and trypsin beads were removed after the centrifugation. 15 µg of protein were desalted on C18 STAGE Tips, dried in rotational vacuum concentrator and dissolved in 25 µl of 0.5 % acetic acid in water (Rappsilber, Ishihama, & Mann, 2003).

LC-MS analysis

5 µL of peptides in 0.5% acetic acid were injected in duplicate on a nano LC-MS/MS system (NanoLC-Ultra, Eksigent, Singapore) coupled to LTQ-Orbitrap Velos mass spectrometer (Thermo Fischer Scientific, USA), using a 240 minutes gradient ranging from 5% to 40% of solvent B (80% acetonitrile, 0.1 % formic acid; solvent A=5 % acetonitrile, 0.1 % formic acid). For the chromatographic separation a 20 cm long capillary (75 µm inner diameter) was packed with 3 µm C18 beads (Reprosil-AQ, Dr. Maisch). On one end of the capillary a nanospray tip was generated using a laser puller (P-2000 Laser Based Micropipette Puller, Sutter Instruments), allowing fretless packing.

The nanospray source was operated with a spray voltage of 1.9 kV and ion transfer tube temperature of 260 °C. Data were acquired in data dependent mode, with one survey MS scan in the Orbitrap mass analyzer (resolution 60000 at m/z 400) followed by up to 20 MS/MS (LTQ-Orbitrap Velos) in the ion trap on the most intense ions (intensity threshold=750 counts). Once selected for fragmentation, ions were excluded from further selection for 30 seconds, in order to increase new sequencing events.

Data processing and analysis

Raw data were analyzed using the MaxQuant proteomics pipeline (v1.5.2.8) and the built-in Andromeda search engine (Cox, et al., 2011) with the updated Uniprot database including isoforms (UniProtconsortium, 2017). Carbamidomethylation of cysteines was chosen as a fixed modification, oxidation of methionine and acetylation of N-terminus were chosen as variable modifications. The search engine peptide assignments were filtered at 1% FDR[1] and the feature match between runs was not enabled; other parameters were left at default settings.

[1] FDR – false discovery rate

After detection and identification, the contaminants (according to the list of contaminants provided by the software), proteins only identified by the modified site, and reversed were removed from the further analysis. The quality control of the proteome data was performed via in-house developed software PTXQC (Bielow, Mastrobuoni, & Kempa, 2016).

Metabolomics

The metabolite extraction procedure consisted of a biphasic solvent extraction consisting of a mixture of methanol-chloroform-water (MCW) (5:2:1 v:v:v) (Methanol LC-MS-grade, Chloroform Reagent Plus ® 99,8% Sigma-Aldrich) with cinnamic acid (2 µg/mL, Sigma-Aldrich) as an internal standard. Samples were lysed in MCW (1 mL per 30 mg of sample) using the tissue lyser (Precellys 24 lysis and homogenization, Bertin Technologies, France) and zirconium beads, samples were cooled on ice between the shaking cycles. Samples were shaken for 60 min at 1000 rpm at 4°C. After addition of ice cold water (at half the MCW volume originally used) samples were shaken for 10 min at 1000 rpm at 4°C. Samples were centrifuged for 10 min at 20000 rcf to separate the polar (top), lipid (bottom) and interphase (tissue debris) layers. The polar phase containing polar metabolites was dried under vacuum for 12 h (in the Rotational vacuum concentrator (RVC) 2-33 CD plus, Christ, Germany).

The derivatization procedure consisted of the dried extracts being dissolved in 10 µL of methoxyamine hydrochloride solution (Sigma, 40 mg/mL in pyridine (Roth)) and incubated for 90 min at 30°C with shaking at 1000 rpm followed by the addition of 40 µl of N-methyl-N-[trimethylsilyl] trifluoroacetamide (MSTFA; Machery-Nagel, Dueren, Germany) and incubation at 37°C for 45 min with shaking at 1000 rpm. The extracts were centrifuged for 10 min at 20000 rcf, and aliquots of 30 µL were transferred into glass vials (Chromacol, UK) for gas chromatography-mass spectrometry (GC-MS) analysis.

The analysis of metabolites was performed with a Pegasus IV mass-spectrometer (LECO, St. Joseph, USA) (details are described in (Pietzke, Zasada, Mudrich, & Kempa, 2014)).

Data analysis: the GC-MS chromatograms were pre-processed with the ChromaTOF software (LECO). Calculation of retention index, mass spectra identification and metabolite quantification were performed using the in house Maui-VIA Software tool (Kuich, et al., 2014). Cinnamic acid was added as an internal control to estimate the extraction quality, and further normalization of the signal intensities. The 75% of cinnamic acid intensity of all samples was used a cut-off line. Samples, which had an intensity of cinnamic acid lower than 75% of the median were excluded from the analysis (one sample in the case of the mouse model).

Visualization of both proteomic and metabolomic data was performed with Vanted software v2.2.1.

Genomics

DNA extraction was performed from biopsy or surgery frozen material with DNeasy Blood and Tissue Kit (catalog number: 69504, Qiagen Inc, USA) according to manufacturer specifications: Purification of Total DNA from Animal Tissues (Spin-Column Protocol) (QIAGEN, 2006). The amount of DNA was measured by NanoDrop (NanoDrop ND-1000, Peqlab, Thermo Fischer Scientific, USA).

Human cohort

All procedures with human samples were performed in accordance to approved ethics protocol EK 206/09.

Material overview

In this study two types of human liver samples were analyzed: bioptate and surgical material. 95 patients were involved in this study. The cohort description and the rationale behind its development is given in the Results section. Amount of material for proteomics (obtained from both bioptate and surgery): 4.0-41.1 mg, for genomics: 6.9-66.0 mg, for metabolomics (the material amount was not sufficient for all the samples due to the small size of biopsy): 7.6-88.0 mg.

The cohort of human serum samples included 5 HCC patients, from whom serum was taken prior to and from 3 to 7 days after liver resection surgery. 5 patients with either hemangioma or cyst formed the control group.

Proteomics

Protein extraction from human samples was performed in the same way as the mice samples as described above.

For all the samples in the time course experiment, 5 µL were injected on a LC-MS/MS system (samples 1-24 [batch 1]: NanoLC 1D Plus (Eksigent, Singapore) coupled to an LTQ-Orbitrap Velos (Thermo Fischer Scientific, USA); samples 25-95 [batch 2], serum samples: QExactive Plus (Thermo Fischer Scientific, USA)), using a 240 minutes gradient ranging from 5% to 45% of solvent B (80% acetonitrile, 0.1% formic acid; solvent A=5% acetonitrile, 0.1% formic acid). For the

chromatographic separation a 20 cm long capillary (75 um inner diameter) was packed with C18 beads (Reprosil-AQ, Dr. Maisch). On one end of the capillary a nanospray tip was generated using a laser puller (P-2000 Laser Based Micropipette Puller, Sutter Instruments).

The nanospray source was operated with a spay voltage of 2.1 kV and an ion transfer tube temperature of 260 degrees. Data were acquired in data dependent mode, with a top 20 method on the LTQ-Orbitrap Velos (batch 1: samples 1-24 of liver biopsies) (one survey MS scan in the Orbitrap mass analyzer, 60000 resolution at 400 m/z, followed by up to 20 MS/MS scans in the ion trap on the most intense ions, intensity threshold=750 counts) or a top 10 method on the QExactive Plus (batch 2: 25-95 of liver biopsies, and serum samples) (one survey MS scan with resolution 70,000 at m/z 200, followed by up to 10 MS/MS scans on the most intense ions, intensity threshold 5,000). Once selected for fragmentation, ions were excluded from further selection for 30 seconds, in order to increase new sequencing events.

Both batches of raw data were analyzed together in one run using the MaxQuant proteomics pipeline (v1.5.3.12) and the built in the Andromeda search engine (Cox & Mann, 2008). MaxQuant enables high peptide identification rates, individualized parts per billion range (ppb-range) mass accuracies and proteome-wide protein quantification (Cox & Mann, 2008) and the updated Uniprot database including isoforms was employed (UniProtconsortium, 2017). Carbamidomethylation of cysteines was chosen as fixed modification, oxidation of methionine and acetylation of N-terminus were chosen as variable modifications. The search engine peptide assignments were filtered at 1% FDR and the feature match between runs was enabled; other parameters were left at their default settings.

After detection and identification, the contaminants (according to the list of contaminants provided by the software), proteins only identified by a modified site, and reversed were removed from the further analysis. The quality control of the proteome data was performed via in-house developed software PTXQC (Bielow, et al., 2016). The 17th sample was excluded from the further data analysis because it did not pass the quality control, by not reaching the cut-off line of 1000 protein identification counts.

The proteomic analysis of human serum samples was performed in the same way.

Metabolomics

The extraction procedure of the metabolites from human samples were performed in the same way as with the mice samples described above. The analysis of metabolites, data analysis

and visualization were performed as described above for the mouse model. The availability of the samples is presented in Appendix 3.

Genomics

The following work was performed in the collaboration with Prof. Dr. Christine Sers and members of her lab: Dr. Soulafa Mamlouk and Andrea Menne. DNA extraction was performed from the biopsy or surgery frozen material with DNeasy Blood and Tissue Kit (catalog number: 69504, Qiagen Inc, USA) according to the following protocol: Purification of Total DNA from Animal Tissues (Spin-Column Protocol) (QIAGEN, 2006). Due to the lack of DNA mutation data on the samples 32, 47, 74, 33, 16 is not available.

Qualification and quantification of the extracted DNA was performed by TaqMan RNase P detection Reagents Kit (catalog number: 4316831, ThermoFisher scientific, USA). The amount of DNA varies from 0.7 to 50.1 ng/µL. The amplification of genomic DNA and the library preparation was performed using an Ion AmpliSeq Cancer HotSpot Panel v2 (CHP primer pool) was used (catalog number: 4475346, ThermoFisher scientific, USA).

Library Preparation was performed in the following way. 10ng DNA was used for PCR amplification. Ion AmpliSeq Library Kit 2.0 (catalog number: 4475346, ThermoFisher scientific, USA) was used: PCR amplification with CHP primer pool; adapter ligation (IonXpress Barcode Adapters 1-16 (catalog number: 4471250, ThermoFisher scientific, USA) and IonXpress Barcode Adapters 17-23 (catalog number: 4474009, ThermoFisher scientific, USA)); purification using Agencourt AMPure beads (catalog number: A63881, Beckman coulter, Germany).

Sequencing of each 32 samples (were pooled in each run) was performed in a single run on an IonTorrent PGM using a 318v2 Chip.

Results

The phenomenon of metabolic reprogramming is one of the hallmarks of cancer. In the current work I aim to evaluate the role of metabolic changes in the hepatocellular carcinoma (HCC) development. The ASV-B mouse model suits this purpose very well: all the mice have the same genetic background, the same known cause of liver cancer and predictable time frames of HCC development. They are also easy to maintain in identical conditions. The mouse model also allows the inclusion of a suitable control group. Metabolomic and proteomic profiles of mice usually show rather stable trends due to the controlled experimental setup. By contrast, humans have heterogeneous genetic backgrounds and live in varying environmental conditions, resulting in many diverse causes (etiologies) of HCC. Numerous associated medical conditions are possible, and several factors, such as dietary habits or alcohol or drug abuse often remain unknown. Thus, analyses of the human proteome and metabolome often reveal less consistent profiles, then in laboratory mice. Researches have to assemble large cohorts to level the great heterogeneity in the human population in order to identify robust changes between HCC and non-HCC conditions in humans. Further comparison of knowledge obtained from both mice and humans would allow identification of universal key features of metabolic changes in HCC.

1. Mice

Mouse model

In the present study I used 7 mice at 16 weeks old with fully developed HCC and 7 C57BL/6J mice of the same age as a healthy control.

Method development

In order to test whether the amount of material taken from each patient would be enough for the planned study, we took HCC material from mice using a biopsy needle. After the collection, the material was immediately frozen in liquid nitrogen and stored at -80 ˚C. Prior to the analysis each sample was manually cut with a scalpel into 3 pieces in a pre-cooled mortar (scalpel, mortar, and forceps were cooled with liquid nitrogen). Biopsy sections were used for the proteomic, metabolomic and genomic analyses (Figure 3). The smaller piece of biopsy was used for proteomic analysis, medium-size part was used for genomic analysis, and the larger part was used for metabolomic analysis. The amount of material used was as follows: for proteomics: 4.4-9.1 mg, for

metabolomics: 3.9-8.9 mg; test extraction of DNA: 4.9-8.2 mg. Further multiomics analysis showed that the amount of material obtained by this procedure is enough for the designed experiments.

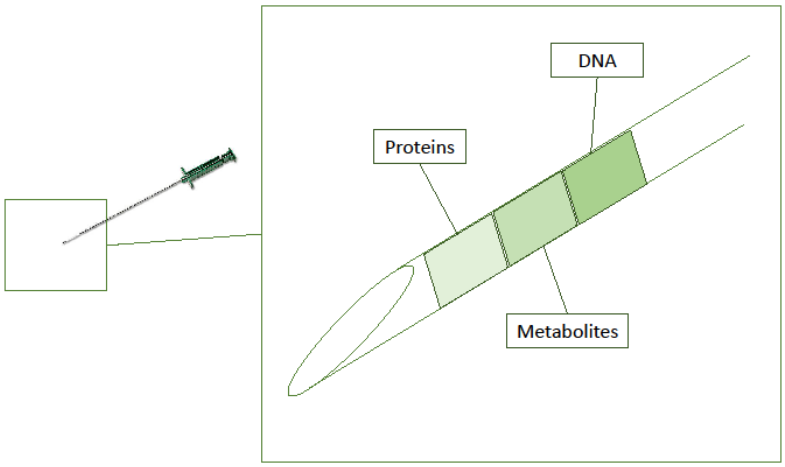


Figure 3. Biopsy sample. Each biopsy was split on 3 parts: for the proteome, metabolome, and genome analyses respectively.

Proteome analysis was performed on the respective part of biopsy. After detection and identification, the contaminants were removed, and thus the output matrix of MaxQuant (maxquant.org) consisted of 4335 proteins. For metabolomics, a GS-MS untargeted analysis of the metabolites extracted from the respective part of biopsy was performed. The investigation was focused on the panel of metabolites, which were detected and identified using the in-house developed software Maui-VIA (Kuich, et al., 2014). Using this developed pipeline and software we detected, identified and quantified 54 metabolites. 11 metabolites which participate in CCM, were used for further analysis. The amount of DNA extracted from the respective part of biopsy was sufficient for the further analysis of equivalent amount of human liver tissue.

Proteomics

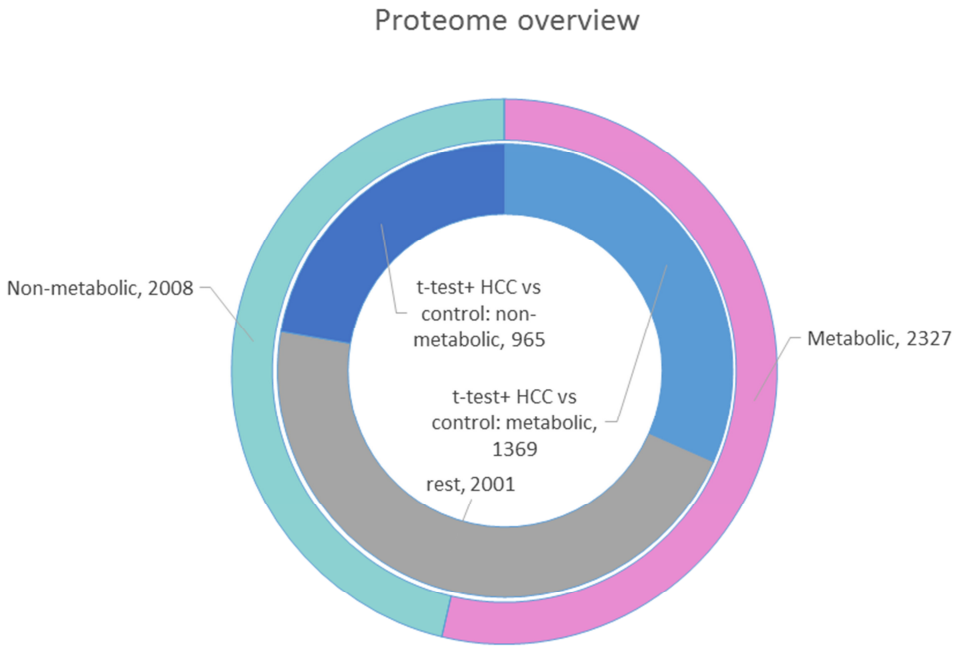
Proteome overview

4335 proteins from 7 tumor and 7 control samples are reported in this study. Gene Ontology (GO) (Ashburner, et al., 2000; GOconsortium, 2015) framework in Perseus (v. 1.5.1.6) to estimate the molecular function of detected proteins was used. According to the GO 53.7% of all proteins detected participate in cellular metabolism. 53.8% of all proteins detected and identified in this experiment were significantly changed according to t-test, and 59% of the significantly changed ones belonged to cellular metabolism (Figure 4A). The t-test with adjustment for multiple testing was performed using the built-in algorithms in Perseus (v. 1.5.1.6). A p-value of below 0.05 was considered significant.

Figure 4B shows the data obtained by different statistical approach, applied in Perseus (v. 1.5.1.6). In order to work with the normally distributed data, it was $\log_2$ transformed. Filtering was applied to each group (control and HCC) to ensure neither group had more than 30% missing values before a two sample t-test was performed with an adjustment for multiple testing. Statistical significance was deemed to be 95% i.e. a p-value of less than 0.05. The output matrix contained 2581 proteins. 59% of fished-out proteins according to GO belonged to the group of cellular metabolism proteins. 72% of proteome was significantly changed in HCC group compared to non-HCC. 60% of significantly changed proteins belonged to cellular metabolism proteins according to GO (t-test, $p<0.05$).

Both applied statistical methods revealed more than 50% of significantly changed proteins.

A.


Proteome overview
Non-metabolic, 2008
t-test+ HCC vs control: non-metabolic, 965
t-test+ HCC vs control: metabolic, 1369
rest, 2001
Metabolic, 2327

B.

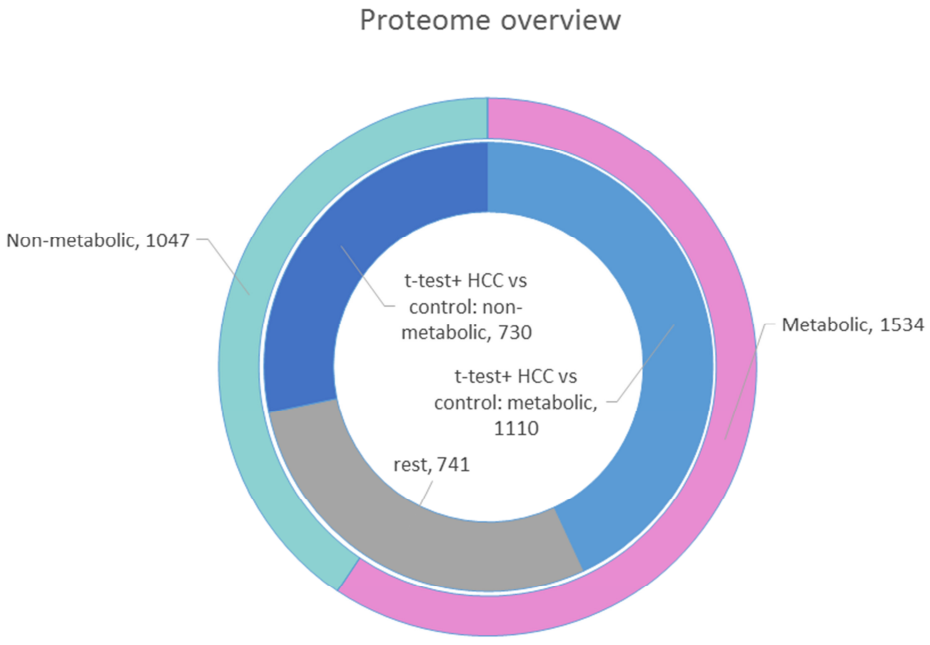

Proteome overview
Non-metabolic, 1047
t-test+ HCC vs control: non-metabolic, 730
t-test+ HCC vs control: metabolic, 1110
rest, 741
Metabolic, 1534

Figure 4. The mice proteome overview. Proteomic data was analyzed using Gene Ontology (GO) (Ashburner, et al., 2000; GOconsortium, 2015) framework in Perseus (v. 1.5.1.6) to estimate the molecular function of detected proteins. **A.** A Student's t-test with adjustment to multiple testing was applied for the MaxQuant output matrix of 4335 proteins. According to the GO, 53.7% of all proteins detected participate in cellular metabolism. 53.8% of all proteins were significantly changed according to Student's t-test, and 59% of the significantly changed proteins were identified by GO as belonging to cellular metabolism (t-test, $p<0.05$). **B** shows the data obtained by different statistical approach, applied in Perseus (v. 1.5.1.6). Log_2 transformed data was used in order to work with the normally distributed data. The filtering on HCC and non-HCC groups were applied for each group to have no more than 30% of missing values, the data imputation (replaced missing values with normal distribution) was applied, and the two sample t-test was applied with the adjustment for multiple testing. The output matrix contained 2581 proteins. 59% of proteins according to GO belonged to the group of cellular metabolism proteins. 72% of proteome was significantly changed in the HCC group compared to control. 60% of significantly changed proteins belonged to cellular metabolism proteins according to GO (t-test, $p<0.05$). Rest – proteins which have not been shown to be significantly changed in the HCC group.

Focused proteome analysis

Further analysis concentrated on the enzymes of CCM and its isoforms. I selected 91 protein which participate in glucose metabolism, TCA cycle, FA synthesis and degradation, glycogen metabolism, pentose phosphate pathway (PPP) (Figure 2). Full names and short names are presented in Appendix 2**.** This list of 91 proteins remains the same for both mouse and human species. However, not all of the listed proteins were detected in the actual proteomic data. Some of proteins were detected only in human (for example, hexokinases). Thus, in the case of the mouse model, this list was shortened to 67 proteins.

Hierarchical clustering in Perseus (v. 1.5.1.6) (parameters: distance - Pearson correlation, linkage – average) was performed on the protein label-free quantification (LFQ) intensities and revealed perfect separation of samples of two distinct groups: tumor and healthy control (**Figure 5**).

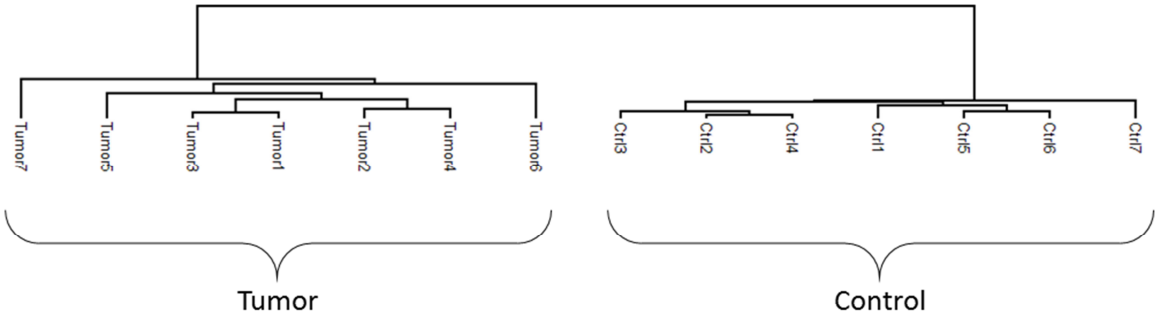


Figure 5. Mice samples clustering based on chosen and detected proteins of CCM. Dendrogram clustering (parameters: distance - Pearson correlation, linkage – average, preprocessed without k-means) performed on LFQ intensities of 67 proteins detected from a starting list of 91 chosen proteins revealed perfect separation of samples of two distinct groups: tumor and healthy control.

Metabolomics

In order to evaluate the changes in CCM at the metabolomic level between tumor and non-tumor conditions, we extracted metabolites from mouse liver material (5 tumor samples and 7 controls). We performed the GC-MS-based untargeted analysis of the metabolites. Further investigation focused on the panel of metabolites, which were detected and identified using the in-house developed software MAUI-VIA (Kuich, et al., 2014; Pietzke, et al., 2014). Using the developed methods and software, we were able to detect and identify 54 metabolites. In the mice samples 11 metabolites, which participate in CCM (Figure 2) were used for the further analysis: glucose, glucoses-6-phosphate, fructose-6-phosphate, phosphoenolpyruvate, lactate, citrate, succinate, fumarate, malate, aspartate, and glutamine.

Multiomics picture

The major aim of this project was to identify differences between HCC and non-HCC metabolism of liver tissue. Vanted software (v. 2.2.1) was used to visualize the relative changes in the main metabolic enzymes at an isoform level, and metabolites. Here and further I compared the changes of mean values of proteins and metabolites intensities.

Proteome and metabolome analyses revealed several striking differences between HCC and normal liver metabolism (Figure 6, 7, 8, 9). Figure 6 represents the overview of CCM changes in

HCC. In liver tissue damaged by HCC the glycogen metabolism was impaired: glycogen synthase (Gys2) level was 3.6 fold lower in HCC compared to control (t-test, $p=3.26*10^{-8}$).

Gluconeogenesis was downregulated at the level of phosphoenolpyruvate carboxykinase 1 (Pck1), which was expressed 16 times less in HCC samples (t-test, $p=2.96*10^{-14}$), G6pc was not detected in HCC samples.

Transport of pyruvate to mitochondria was impaired at the level of mitochondrial pyruvate carrier 1 (Mpc1): 4.2 times lower in HCC (t-test, $p=1.08*10^{-5}$), and Mpc2: 1.3 times lower[1] in HCC (t-test, $p=2.54*10^{-3}$) (Figure 7).

60% of TCA cycle enzymes were downregulated, for example citrate synthase (Cs) expression was 2.6 times lower in HCC (t-test, $p=2.18*10^{-14}$), and succinate dehydrogenases A and B (Sdha and Sdhb) were expressed 2 and 3 times less respectively in HCC (t-test, $p=1.14*10^{-10}$ and $p=2.96*10^{-14}$ $p=3.58*10^{-16}$ respectively) (Figure 8).

Enzymes, which are responsible for the glutamine uptake to the TCA cycle showed lower expression in the HCC samples. Glutamate pyruvate transaminase (Gpt), glutamate pyruvate transaminase 2 (Gpt2), and glutaminase 2 (Gls2) were 1.8, 1.5, and 11 times lower in HCC, respectively (t-test, $p=5.51*10^{-8}$, $p=3.71*10^{-5}$, $p=6.83*10^{-14}$) (Figure 8).

Phosphoglycerate mutases (Pgm) and pyruvate kinases (Pk) showed isoform switches, which represents one of most striking features of metabolic changes in cancer.

At the metabolomic level the upregulation of glycolysis could be observed: glucose-6-phosphate level in the tumor was 3.2 times higher (t-test, p=0.0006), fructose-6-phosphate 3.2 times higher in tumor, (t-test, p=0.0007). Lactic acid level was 1.4 times higher (t-test, p=0.007), fumaric acid 1.4 times higher in HCC samples (t-test, p=0.03), and malic was 1.9 times higher in HCC (t-test, $p=5.16*10^{-7}$) (Figure 9) compared to control.

[1] Here and further it is important to note, that while statistically speaking this difference is significant, its biological impact is questionable. However, even a small fold change in a hormone or in a key regulatory enzyme may have large consequences. Thus, I chose to present the data even though the fold change is less than 2, because the trends of changes are relevant in this work.

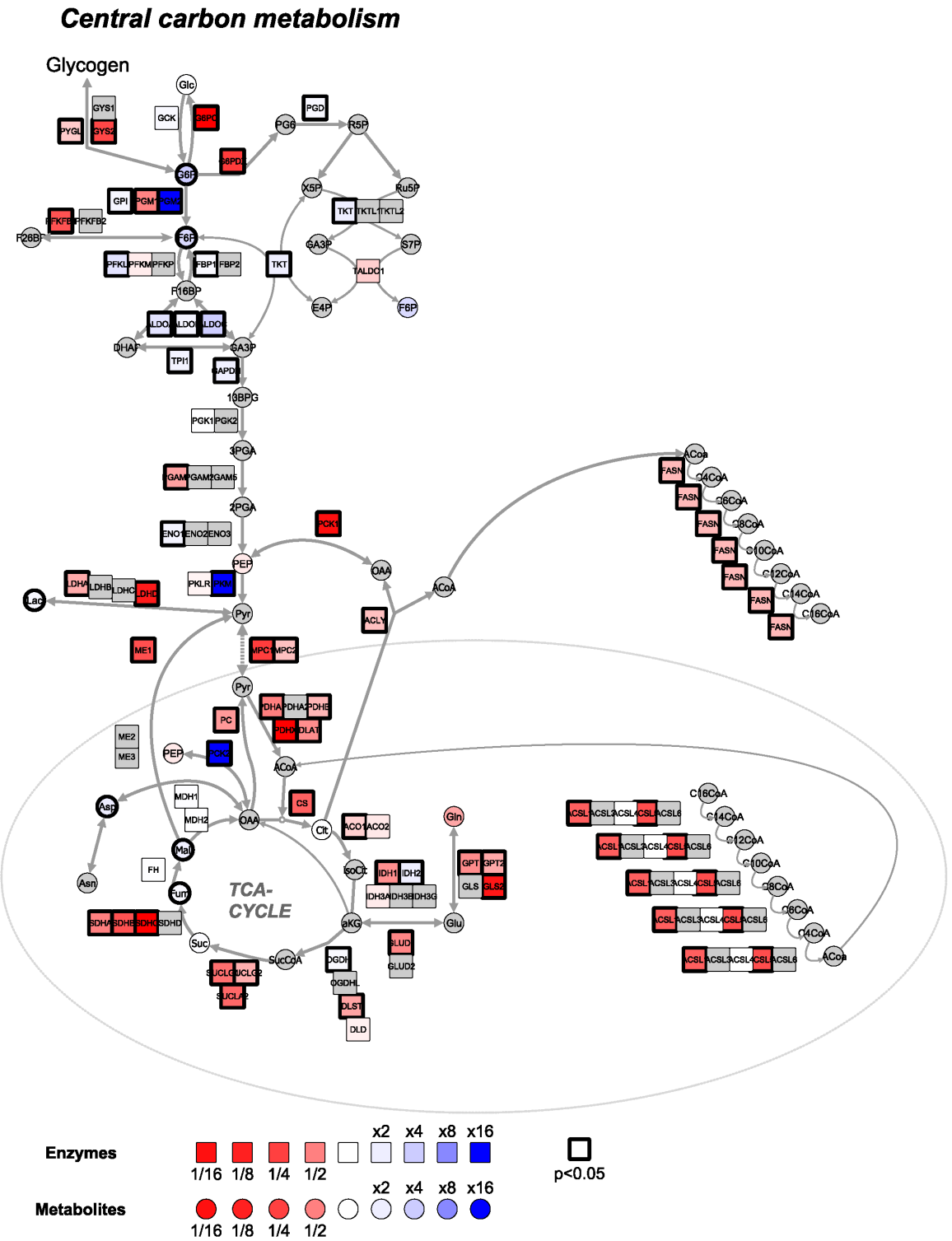


Figure 6. Changes in CCM in mouse HCC tissue compared to a control group. The figure visualizes with a colour code the changes of metabolism in liver tissue damaged by HCC compared to a group of control samples at the proteomic and metabolomic levels. The glycogen metabolism at the proteomic level was perturbed at the level of Gys2 enzyme (here and further proteins are

visualized with squares): it was 3.6 fold lower in HCC compared to a control group (t-test, $p=3.26*10^{-8}$). Gluconeogenetic enzyme Pck1 was expressed 16 times less in HCC samples compared to a control group (t-test, $p=2.96*10^{-14}$), while G6pc enzyme was not detected in HCC samples (and detected in control samples). Mpc1 (responsible for the transport of pyruvate to mitochondria) was expressed 4.2 times lower in HCC compared to a control group (t-test, $p=1.08*10^{-5}$), and Mpc2 (participates in the transport of pyruvate to mitochondria as well) was expressed 1.3 times lower in HCC (t-test, $p=2.54*10^{-3}$). Approximately 60% of TCA cycle enzymes were downregulated. Cs enzyme was expressed 2.6 times lower in HCC (t-test, $p=2.18*10^{-14}$), and Sdha and Sdhb were expressed 2 and 3 times lower in HCC (t-test, $p=1.14*10^{-10}$ and $p=3.58*10^{-16}$ respectively). Sdhc was not detected in HCC samples, while it was detected in the samples of a control group. Other TCA cycle enzymes, such as Gpt, Gpt2, Gls2 were 1.8, 1.5, and 11 times lower in HCC, respectively (t-test, $p=5.51*10^{-8}$, $p=3.71*10^{-5}$, $p=6.83*10^{-14}$). Isoform switch in HCC samples was detected at the level of two glycolytic enzymes: PGMs, PKs. Changes in metabolites (visualized with circles) were detected in glycolysis. Glucose-6-phosphate level was 3.2 times higher in HCC samples compared to a control group (t-test, p=0.0006), fructose-6-phosphate is 3.2 times higher in HCC samples compared to a control group, (t-test, p=0.0007). Several metabolites of TCA cycle undergo changes. The detected lactic acid level was 1.4 times higher (t-test, p=0.007) in HCC samples compared to a control group, fumaric acid was 1.4 times higher in HCC samples (t-test, p=0.03) compared to a control group, and malic acid was 1.9 times higher in HCC group of samples (t-test, $p=5.16*10^{-7}$) compared to a control group. The metabolites are visualized with circles, the proteins are visualized with squares, bold frame indicates the changes with $p<0.05$ (t-test), grey colour of the metabolite (circle) or protein (square) at the scheme indicates that data regarding certain enzyme or metabolite was not available. Red colour indicates decrease of intensity in HCC samples, blue colour indicates increase of intensity in HCC samples.

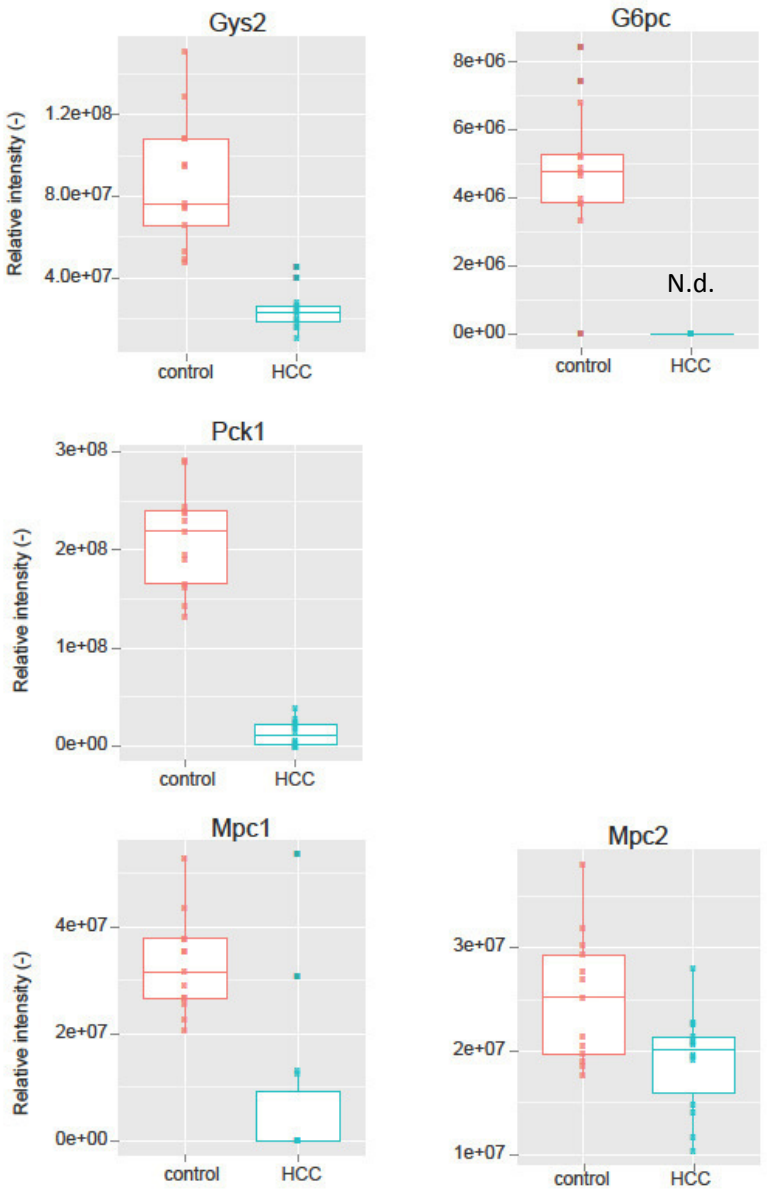


Figure 7. Box plots with the relative intensities of glycolytic and gluconeogenetic enzymes, and mitochondrion pyruvate transporters in mouse HCC samples compared to a control group. An intensity of Gys2 enzyme was 3.6 lower in HCC samples compared to a control group (t-test, $p=3.26*10^{-8}$), G6pc was not detected in HCC samples while detected in the samples of control group, Pck1 intensity was 16 times lower in HCC samples compared to a control group (t-test, $p=2.96*10^{-14}$). Mpc1 intensity was 4.2 times lower in HCC group of samples compared to a control group (t-test, $p=1.08*10^{-5}$), and Mpc2 intensity was 1.3 times lower in HCC samples compared to control samples (t-test, $p=2.54*10^{-3}$). N.d. – not detected.

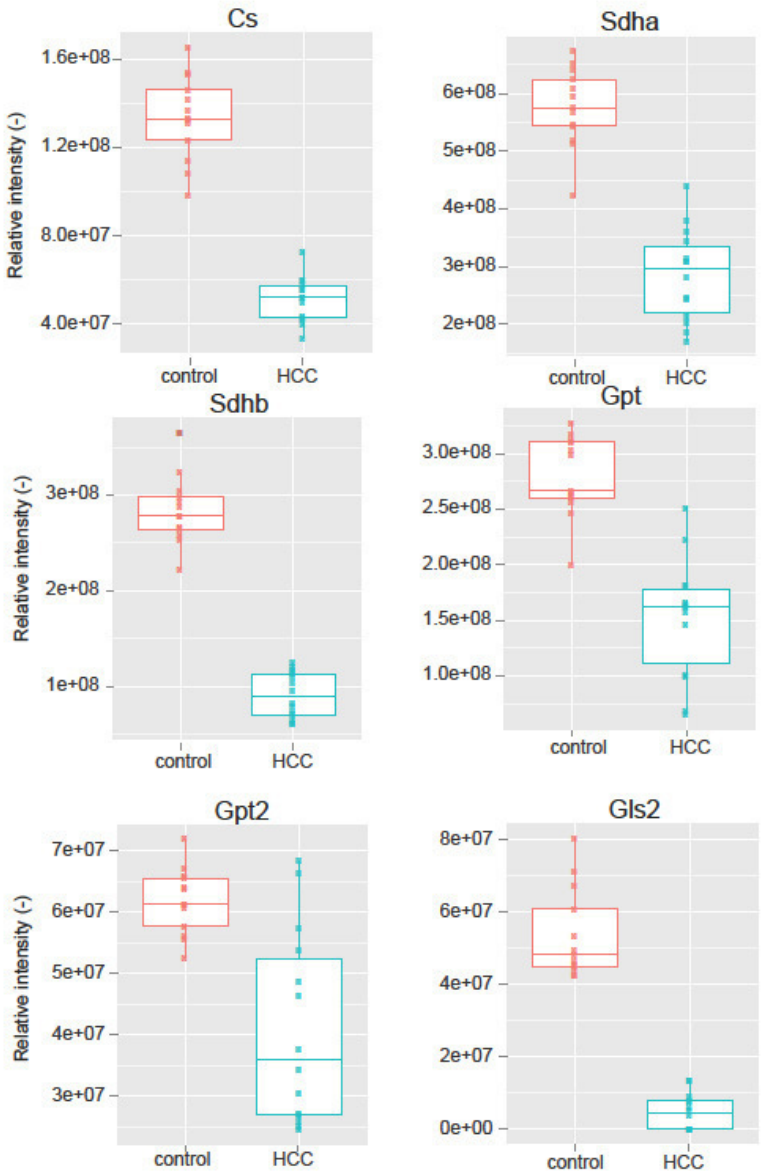


Figure 8. Box plots with the relative intensities of TCA cycle and glutaminolysis enzymes in mouse HCC samples compared to a control group. Relative intensity of Cs enzyme was 2.6 times lower (t-test, $p=2.18*10^{-14}$) in HCC samples compared to a control group, an intensity of Sdha was 2 times lower in HCC samples compared to a control group (t-test, $p=1.14*10^{-10}$), an intensity of Sdhb was 3 times lower (t-test, $p=3.58*10^{-16}$) in HCC samples compared to a control group.

Relative intensities of Gpt, Gpt2, Gls2 were 1.8, 1.5, and 11 times lower in HCC compared to a control group, respectively (t-test, $p=5.51*10^{-8}$, $p=3.71*10^{-5}$, $p=6.83*10^{-14}$).

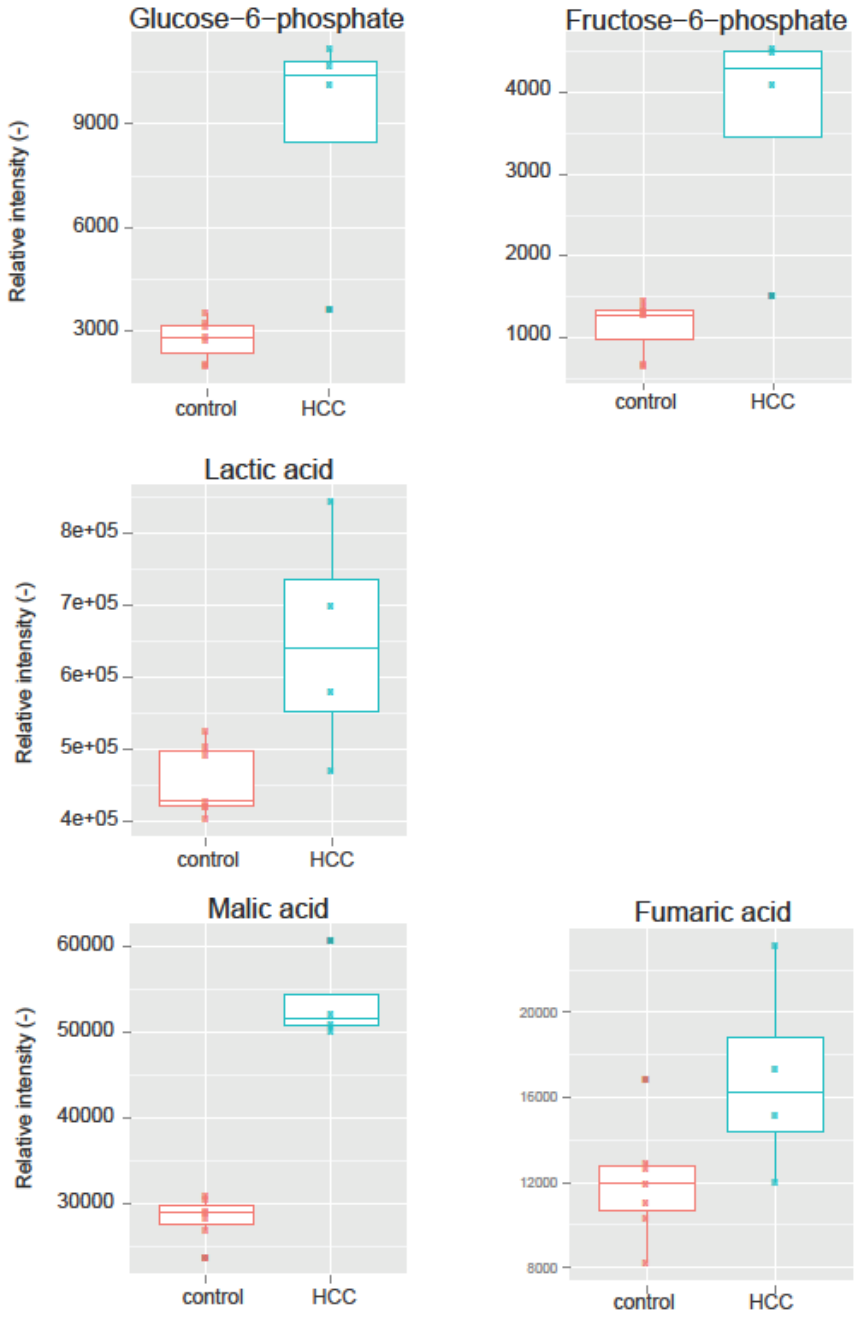


Figure 9. Box plots with the relative intensities of detected metabolites of CCM in mouse HCC compared to a control group. Relative intensity of glucose-6-phosphate was 3.2 times higher in HCC samples compared to a control group (t-test, p=0.0006), fructose-6-phosphate intensity was

3.2 times higher in HCC samples compared to a control group, (t-test, p=0.0007). Relative intensities of TCA cycle enzymes were the following: lactic acid intensity was 1.4 times higher in HCC samples compared to a control group (t-test, p=0.007), an intensity of fumaric acid was 1.4 times higher in HCC samples in HCC samples compared to a control group (t-test, p=0.03), and malic acid intensity was 1.9 times higher in HCC compared to a control group (t-test, $p=5.16*10^{-7}$).

2. Human cohort

Formation of the patient cohort

This study focuses on the comparison of the main metabolic changes between liver samples with HCC and samples with pre-cancerous conditions and comparatively mild liver conditions, such as fatty liver. Several possible confounding factors had to be considered during the cohort formation to account for the great heterogeneity among patients. Nutritional factors/diet and alcohol abuse are usually difficult to track, due to the incomplete anamnesis collection and patients not willing to admit to any illegal actions. We excluded patients (according to their anamnesis) with known exposure to certain hepatotoxins (e.g. aflatoxin, industrial toxins), as well as patients with long-term usage of certain medication (6 month of use or longer) (for example, insulin or aspirin) (to avoid including various drug-induced liver diseases in the cohort). Patients with hereditary conditions, metabolic disorders or autoimmune diseases were excluded from the second part of cohort (samples 36-95). Thus the patients with hemochromatosis, Gilbert's syndrome, Dubin–Johnson syndrome, Crigler–Najjar syndrome, Wilson's disease, alpha 1-antitrypsin deficiency, mucoviscidosis, galactosemia, dyslepidemia, and others were not included in the cohort. However, a few samples from patients with diabetes, obesity, and alcohol addiction happen to be in the cohort[1] (for example, the information about BMI came to the lab later than the actual sample).

Current hepatitis viral load was considered, and a separate groups of patients with hepatitis virus B or C were formed. Unfortunately, we were not able to track whether the patients had a history of previous hepatitis exposure. Patients with other detected viruses (e.g. HIV) were excluded from the cohort. We also excluded patients who required regular blood transfusions, because the iron overload of the liver could affect the metabolic picture. Liver cancer types were histologically diagnosed in the clinic. All HCC samples were diagnosed with primary tumors. To the

[1] Diabetes type 1: 21; diabetes type 2:19, 23, 24, 28; patients 18, 20, 24, 46, 49, 56, 65 suffered from obesity (BMI ≥30 kg/m^2, recidive HCC: 74, 84; alcohol addiction: 16, 17, 18, 23, 26, 30, 64, 80, and 93.

best of our knowledge patients did not get any treatment (radionuclide treatment, percutaneous ethanol injections, chemotherapy etc.) prior to the surgery or the biopsy procedures. Blood tests (aspartate transaminase (AST), alanine transaminase (ALT), AFP, blood count, biochemistry) are the subject of further investigation. Figure 10 shows the heterogeneity of the diseases and conditions of the patients used in this study. Full diagnosis and additional information including the availability of material for full multiomics analysis is to be found in Appendix 3.

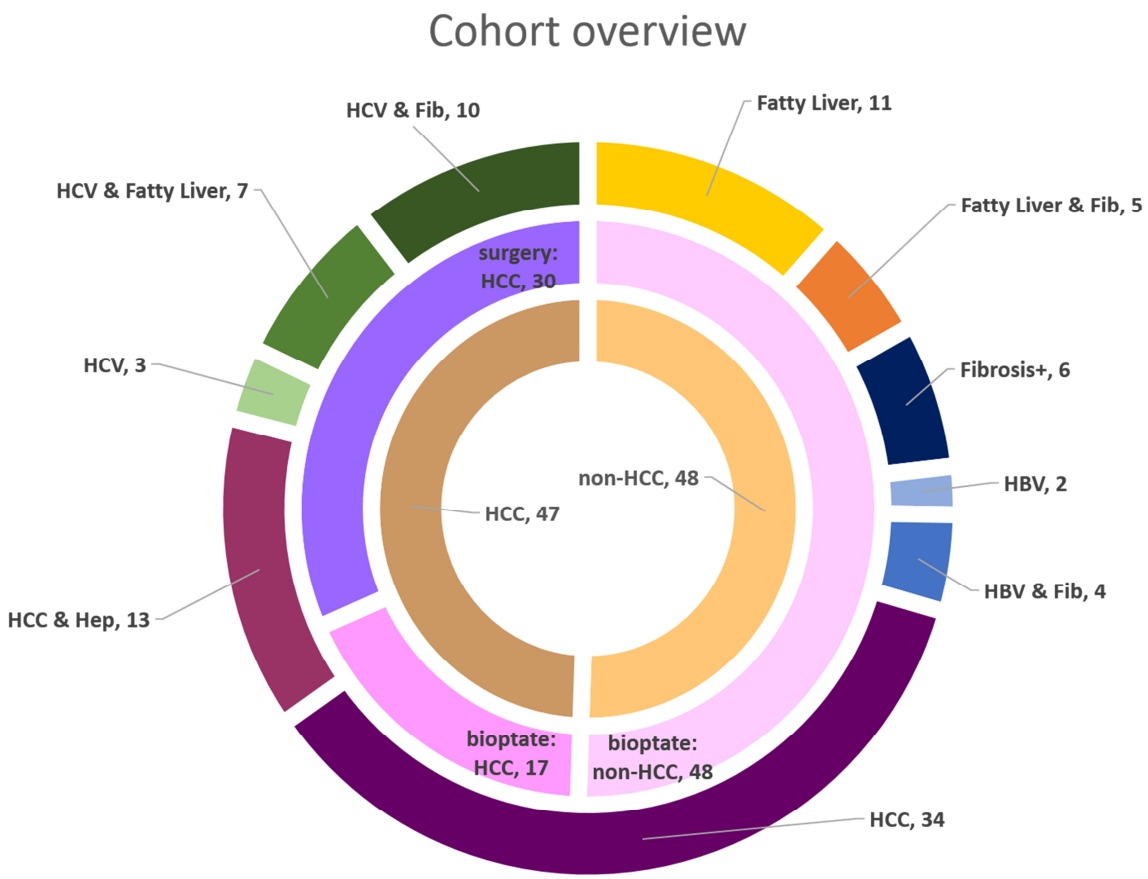


Figure 10. Human cohort overview. The cohort consisted of 95 patients with various conditions, such as fatty liver, fibrosis, cirrhosis, hepatitis B and C, and HCC.

Human liver material overview

In this study two types of human liver samples were analyzed: bioptate and surgery material. The human biopsy material was taken under the control of CT from the patients prior to the treatment. Since we were interested in the investigation of the *in situ* picture, this study was focused only on patients prior to any medical intervention (radionuclide treatment, percutaneous ethanol injections, chemotherapy etc.). The surgical material was obtained during the resection of liver damaged by HCC.

A total of 6213 proteins were detected in the 94 samples (excluding the 17th) and included in the study. Using the developed pipeline and the software we were able to detect and identify 89 metabolites, 8 of these which participate in CCM, were used for further analysis. Our cohort included 95 patients, however the amount of material for targeted genomic analysis was only available from 90 patients due to the lack of material. Mutations in 10 different genes were detected in the cohort.

Proteomics

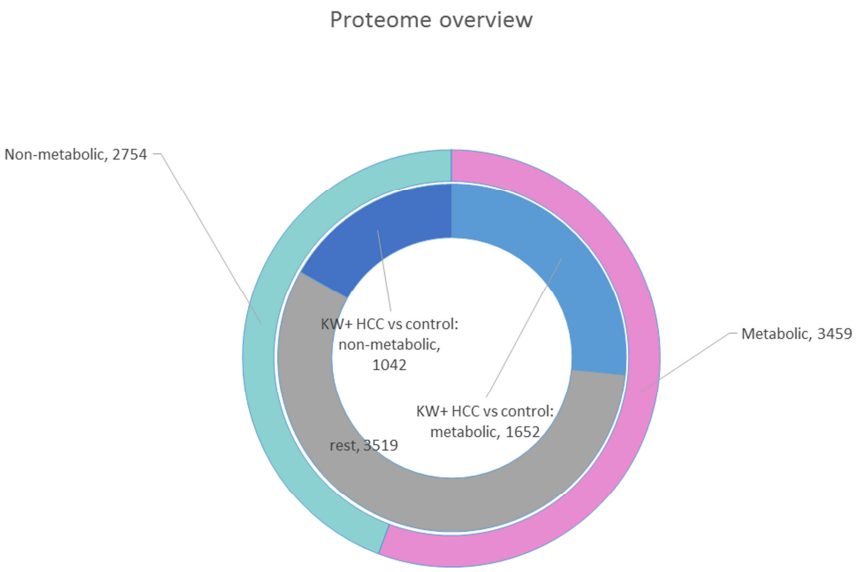
Proteome overview

The proteome analysis was performed in Perseus (v. 1.5.1.6). As shown in Figure 11A, 56% of proteins according to GO belonged to the group of cellular metabolism proteins. A Kruskal-Wallis-Test (KW-test) (the choice of method was performed according to the article published by du Prel (du Prel, Röhrig, Hommel, & Blettner, 2010)) with adjustment for multiple testing revealed that 44% of proteins were significantly changed in tumor samples compared to control, 61% of which belonged to metabolic proteins (KW-test, $p<0.05$).

Figure 11B shows the data obtained by a different statistical approach, applied in Perseus (v. 1.5.1.6). In order to work with normally distributed data, it was $\log_2$ transformed. The filtering for HCC and non-HCC groups was applied such that each group has no more than 30% missing values, the two sample t-test was applied with the permutation-based FDR and adjustment for multiple testing was performed (with the Perseus build-in algorithms). The output matrix contained 2556 proteins. 67% of proteins according to GO belonged to the group of cellular metabolism proteins. 65% of proteome was significantly changed in HCC group compared to non-HCC. 64% of significantly changed proteins belonged to cellular metabolism proteins according to GO (t-test, $p<0.05$).

Both applied statistical methods (KW and t-test) revealed approximately 50% of significantly changed data. Reasons for such high number of significantly changed proteins are reviewed in the Discussion section.

A.



B.

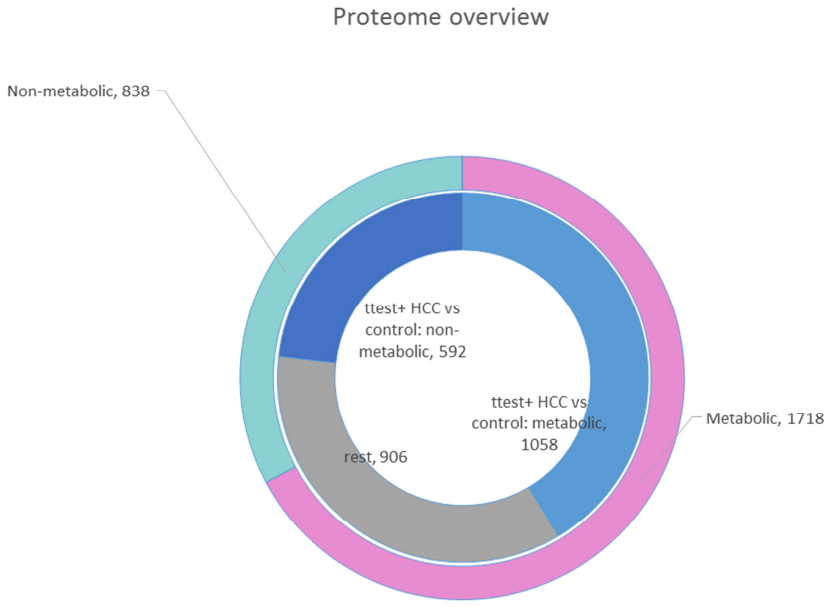


Figure 11. Human proteome overview. A. A total of 6213 proteins were annotated with Gene Ontology (GO). According to GO, 56% of proteins belonged to the group of cellular metabolism proteins. A Kruskal-Wallis test with correction for multiple testing revealed expression of 44% of proteins as significantly changing in tumor samples compared to control, 61% of which belonged to the class of metabolic proteins (KW-test, $p<0.05$). **B.** In order to work with normally distributed data, data was $\log_2$ transformed. Filtering was applied to both HCC and non-HCC groups such that neither group has more than 30% missing values. Following this, a two sample t-test was applied to the data with a permutation-based FDR and adjustment for multiple testing. The output matrix contained 2556 proteins. 67% of these proteins (according to GO) belonged to the group of cellular metabolism proteins. 65% of the proteome was significantly changed in the HCC group compared to non-HCC. 64% of the significantly changed proteins belonged to the group of cellular metabolism proteins according to GO (t-test, $p<0.05$). Rest refers to the proteins where no statistically significant change was observed.

Focused proteome analysis

The subsequent analysis was focused on CCM enzymes and enzymes and their isoforms which are responsible for the main liver metabolic functions. The same 91 proteins were used which were chosen for the mouse model analysis (Figure 2**,** Appendix 2). However, in the case of human samples, not all of the listed proteins were detected in the obtained proteomics data. Thus, this list was shortened to 87 proteins.

Additionally, in Perseus (v. 1.5.1.6), I performed a hierarchical clustering analysis based on proteomic data (parameters: distance - Pearson correlation, linkage – average), displayed in Figure 12**.** This figure displays the clustering according to the label-free quantification (LFQ) values of the 91 proteins mentioned above in patients 01-95. Pearson correlation was selected as a similarity measure since it ensures internal normalization in contrast to Euclidian distance. I also performed the clustering using Spearman correlation, which showed similar trends (Appendix 4). Since the calculation of the Spearman correlation requires an additional ranking step and reduces the information of the data, we decided to work further with the Pearson correlation. However both Euclidian and Spearman clustering show similar trends for the separation of different groups.

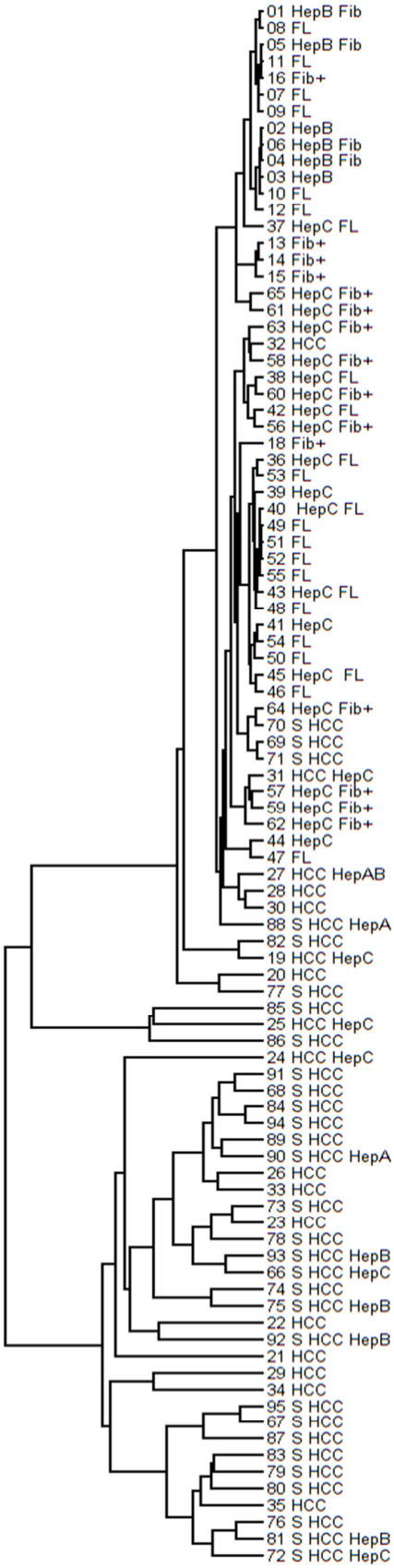

01 HepB Fib
08 FL
05 HepB Fib
11 FL
16 Fib+
07 FL
09 FL
02 HepB
06 HepB Fib
04 HepB Fib
03 HepB
10 FL
12 FL
37 HepC FL
13 Fib+
14 Fib+
15 Fib+
65 HepC Fib+
61 HepC Fib+
63 HepC Fib+
32 HCC
58 HepC Fib+
38 HepC FL
60 HepC Fib+
42 HepC FL
56 HepC Fib+
18 Fib+
36 HepC FL
53 FL
39 HepC
40 HepC FL
49 FL
51 FL
52 FL
55 FL
43 HepC FL
48 FL
41 HepC
54 FL
50 FL
45 HepC FL
46 FL
64 HepC Fib+
70 S HCC
69 S HCC
71 S HCC
31 HCC HepC
57 HepC Fib+
59 HepC Fib+
62 HepC Fib+
44 HepC
47 FL
27 HCC HepAB
28 HCC
30 HCC
88 S HCC HepA
82 S HCC
19 HCC HepC
20 HCC
77 S HCC
85 S HCC
25 HCC HepC
86 S HCC
24 HCC HepC
91 S HCC
68 S HCC
84 S HCC
94 S HCC
89 S HCC
90 S HCC HepA
26 HCC
33 HCC
73 S HCC
23 HCC
78 S HCC
93 S HCC HepB
66 S HCC HepC
74 S HCC
75 S HCC HepB
22 HCC
92 S HCC HepB
21 HCC
29 HCC
34 HCC
95 S HCC
67 S HCC
87 S HCC
83 S HCC
79 S HCC
80 S HCC
35 HCC
76 S HCC
81 S HCC HepB
72 S HCC HepC

Figure 12. Human samples clustering based on chosen and detected 87 protein of CCM. Clustering according to the LFQ intensities of 91 proteins of patients 01-95 (the 17th sample was excluded, because it did not pass the quality control (the explanation in the text)). Parameters: distance - Pearson correlation, linkage – average.

Metabolomics

In order to evaluate the changes of CCM at the metabolomic level for tumor and non-tumor conditions, I extracted metabolites from human liver material from clinical cases (see the availability of the material in Appendix 3). I performed GC-MS untargeted analysis of metabolites. The further investigation was focused on the panel of metabolites which were detected and identified using the in-house developed software Maui-VIA (Kuich, et al., 2014). Using the developed pipeline and the software we could detect and identify 89 metabolites. In the experiment with human samples, 8 metabolites which participate in CCM (Figure 2) were used for the further analysis: lactate, pyruvic acid, citrate, glutamic acid, succinic acid, fumaric acid, malic acid, and aspartate.

It is crucial to underline that in general an overall increase in metabolite signal in the samples from the surgery was detected compared to the signal obtained from the biopsy material. Figure 13 shows the intensities of pyruvic, succinic, and citric acids in two groups of samples. The upper panel contains groups of samples with both surgical and biopsy material, the lower panel shows only the group without surgically delivered samples (only biopsies). All metabolites shown in the upper panel showed higher average intensities in the HCC group compared to tumor-free samples. Out of all detected and identified metabolites only 5% were higher in the tumor-free group of samples. The lower panel shows the intensities of metabolites extracted only from biopsy material. Out of all metabolites identified from biopsy material, 56% were at higher intensities in the tumor-free group. This unequal distribution of intensities of metabolites could be explained by the surgical conditions. The conditions under which the material from the patient is obtained takes several hours due to the necessary actions, such as fasting of the patient prior to the surgery, anesthesia, medication, clipping of the arteries and veins of the liver prior to the procedure etc. This leads to the adjusted metabolome profile, possibly determined by the hypoxic environment. We propose that the metabolome data obtained from the surgical material does not represent the original picture *in situ* of the intact liver. Thus, for the further metabolome analysis we decided to use only the data obtained from the bioptate material.

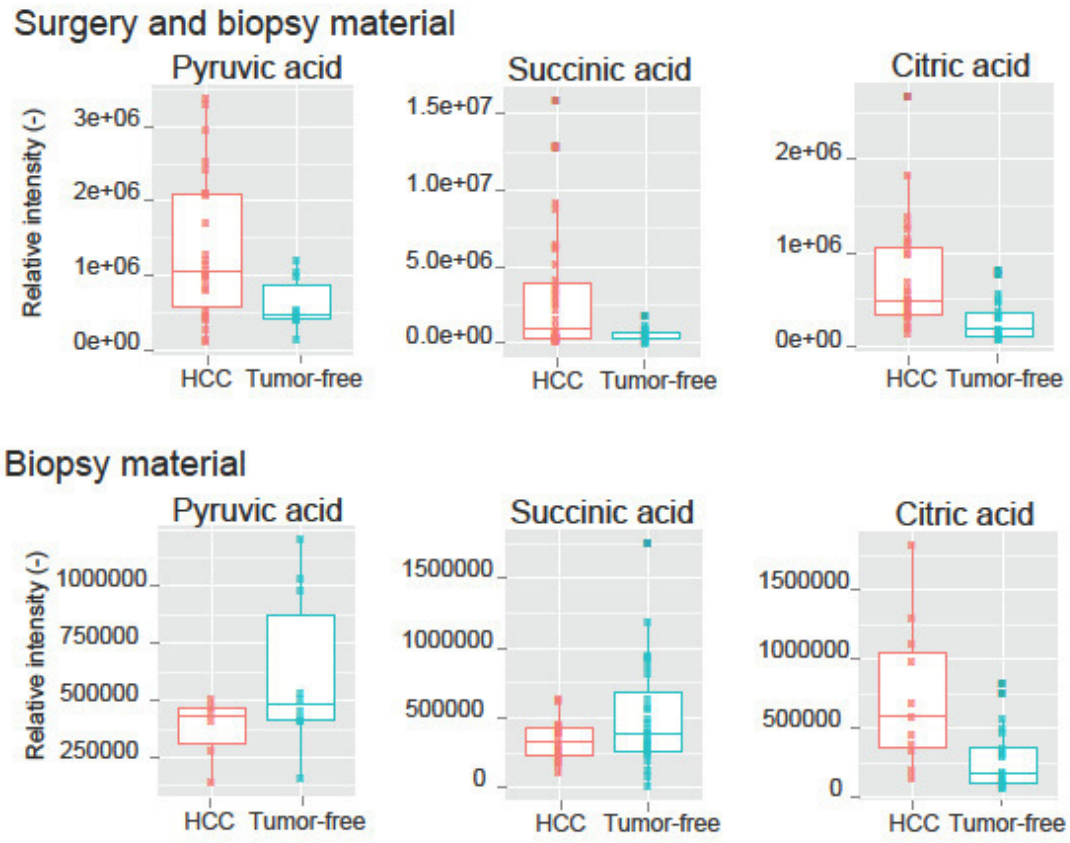


Figure 13. Box plots of recorded relative intensities of selected metabolites from surgical and biopsy material in HCC samples and tumor-free samples. Relative intensities of pyruvic, succinic, and citric acids are shown as two groups of samples. In the upper panel: values for pyruvic acid, succinic acid and citric acid obtained from the samples collected from both surgery and biopsy material. The average intensities of the measured metabolites in the HCC group were higher than in the tumor-free group. Lower panel: values for pyruvic acid, succinic acid and citric acid obtained from the biopsy samples exclusively showed different distribution of metabolites intensities, as compared to the values, obtained from surgery material.

Genomics

Our cohort included 95 patients, however, due to restrictions in amount of available material, only 90 patients were sequenced (excluding #16, 32, 33, 47, and 74 due to lack of

material). Sequence reads were aligned to the GRCh37 sequence using the Torrent Mapping and Alignment Program (TMAP) (ThermoFisher Scientific, USA). Variants were called on the processed reads using the Torrent Variant Caller (TVC) (ThermoFisher Scientific, USA) under the "strict" setting as specified by the IonTorrent Suite. The resulting data was analyzed with an in-house pipeline (developed by our collaborators in the lab of Prof. Dr. rer. nat. Christine Sers). We used a cut-off of 4% allele frequency (Table 2). Further analysis included removing known SNPs (The Genomes Project, 2015) and low quality reads (a threshold of quality Q=50 was used). Questionable mutations were verified by Sanger sequencing (CTNNB1 mutation in sample 95 was confirmed, SMAR mutation in sample 82 was found to be an error). Table 2 includes the verified mutations list: the mutated gene, exact mutation, patient code and diagnosis.

The most commonly mutated gene in the cohort was TP53 gene (approximately 40% of all mutations detected): TP53 mutations were detected 13 times in the cohort (15% of all sequenced patients). 60% of TP53 mutations were detected in HCC samples, 40% in non-HCC samples. The second most commonly mutated gene is CTNNB1: 8% of sequenced patients had mutation in this gene (21% of all mutations in the cohort, detected only in HCC samples). Other genes (APC, ATM, CDKN2A, NOTCH1, PIK3CA, PTEN, RB1, STK11) were found to be mutated less frequently, in both HCC and non-HCC samples. No clear distinction between HCC and non-HCC samples was possible to perform exclusively from targeted mutation analysis.

Table 2. Detected, identified and verified somatic tumorigenic mutations in human cohort. Additional information about patients can be found in Appendix 3. HAV, Hepatitis A virus; HBV, Hepatitis B virus; HCV, Hepatitis C virus; Fibrosis+, Advanced Fibrosis/Cirrhosis; Diab2, Diabetes Type 2;, del, deletion; ins, insertion; fs, frame shift.

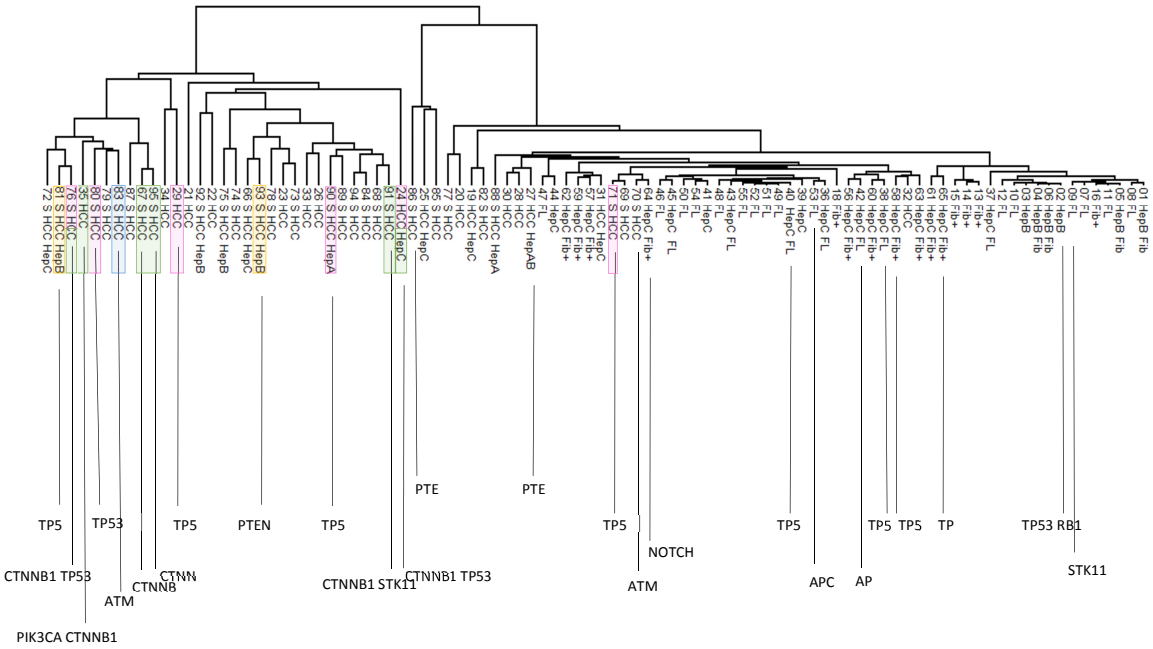
Gene	Mutation	Amino acid	Diagnosis	Patient ID	Allele frequency
TP53	c216delC	p.V73fs	OGDH	2	0.32
RB1	c.C1966T	p.R656W	HBV	2	0.44
STK11	c.C1062G	p.F354L	Fatty Liver	9	0.48
CTNNB1	c.A121G	p.T41A	HCC, HCV, Diab2, Obesity	24	0.19
TP53	c544_550del	p.182_394del	HCC, HCV, Diab2, Obesity	24	0.22
PTEN	c.T882G	p.S294R	HCC, HAV, HBV	27	0.08
TP53	c.C452A	p.P151H	HCC	29	0.81
PIK3CA	c.A3140G	p.H1047R	HCC	35	0.09
CTNNB1	c.T133C	p.S45P	HCC	35	0.22
CDKN2A	c.A248G	p.H83R	HCC	35	0.21
TP53	c216delC	p.V73fs	HCV, Fatty Liver	38	0.82
TP53	c216delC	p.V73fs	HCV, Fatty Liver	40	0.80
APC	c3916_3917insA	p.I1307fs	HCV, Fatty Liver	42	0.08
APC	c3916_3917insA	p.I1307fs	Fatty Liver	53	0.09
TP53	c216delC	p.V73fs	HCV, Fibrosis+	58	0.80
NOTCH1	c4732_4734del	p.1578_1579del	HCV, Fibrosis+, Alcohol	64	0.04
TP53	c.A704G	p.N235S	HCV, Fibrosis+, Obesity	65	0.51
CTNNB1	c.C110T	p.S37F	HCC	67	0.23
ATM	c.T2572C	p.F858L	HCC	70	1.00
TP53	c216delC	p.V73fs	HCC	71	0.79
CTNNB1	c.C134T	p.S45F	HCC	76	0.42
TP53	c216delC	p.V73fs	HCC	76	0.72
TP53	c902delC	p.P301fs	HCC, Alcohol	80	0.62
TP53	c.T581G	p.L194R	HCC, HBV	81	0.05
ATM	c.A8626G	p.I2876V	HCC	83	0.54
CTNNB1	c.C134T	p.S45F	HCC	83	0.56
PTEN	c.G754T	p.D252Y	HCC	86	0.64
TP53	c.C298T	p.Q100X	HCC, HAV, HBV	90	0.07
STK11	c.C1040G	p.A347G	HCC	91	0.51
CTNNB1	c.C134T	p.S45F	HCC	91	0.07
PTEN	c.A203G	p.Y68C	HCC, HBV, Alcohol	93	0.58
TP53	c902delC	p.P301fs	HCC, HBV, Alcohol	93	0.62
CTNNB1	c.C98T	p.S33F	HCC	95	0.39

Multiomics picture

Classification

Classification based on mutations

Several possibilities of classification of HCC are currently discussed in the scientific community. We decided to start our work from the one suggested by Boyault et al. in 2007 (Boyault, et al., 2007) and adapted by Beyoğlu et al. in 2013 (Beyoglu, et al., 2013). We tested whether this classification is applicable for the cohort we work with. Figure 14 displays the clustering according to the LFQ intensities of the 87 above mentioned proteins of samples 01-95 (except 17). Mutations found are indicated with the name of the mutated gene in the lower part of the figure. Different colours represent different groups of HCC, according to the classification of Boyault (Boyault, et al., 2007). The primary decision on the grouping was based on the mutation status of the HCC samples. Since both G5 and G6 groups include CTNNB1 mutations, further distinction was based on e-cadherin expression. Samples which are not labeled with any colour do not fit in the any of the suggested groups. Samples labeled with two colours fit in both groups of HCC. Less than 1/3 of all HCC samples could be allocated to a group according to the features used by this classification system. Groups G1 and G4 did not include any human sample from the cohort I worked with. It is important to note that mutations were not detected in all of the samples. Various mutations existed in both HCC groups, and non-cancerous conditions, the lesion group also included patients where no mutations were found.



Transcriptomic group	Characteristics
G1	Associated with HBV infection Low copy number of viral DNA *AXIN1* mutations Younger age High serum AFP Frequently of African origin Overexpressed genes: *MYH4, SOX9, IGF2, PEG3, PEG10, AFP, SGCE*
G2	Associated with HBV infection High copy number of viral DNA Frequent local and vascular invasion *TP53* and *AXIN1* mutations
G3	*TP53* mutations without HBV infection Chromosome losses at 17p, 5q, 21q, and 22q *CDKN2A* ($p16^{INK4a}$ gene) promoter hypermethylation Overexpressed genes: *CDC6, MAD2L1, BUB1, TTK, SMC1L1, CCNA2, CCNE2, MCM2, MCM3, MCM6, ASK, KPNB1, RANBP5, XPO1, IPO7, NUP155, NUP107*
G4	Heterogeneous group of tumours that clustered with nontumorous livers 4/20 had *TCF1* (HNF1A gene) mutation and clustered with hepatocellular adenoma and well-differentiated HCC
G5	Activation of Wnt and β-catenin (70% *CTNNB1* mutations) Overexpressed genes: *GPR49, GLUL, PAP/HIP* (β-catenin target genes in the liver) Underexpressed genes: *ARHG-DIB, HLA-DPA1/B1, IFI16, IFI44, PTGER4, STAT1, CLECSF2*
G6	Activation of Wnt and β-catenin (100% *CTNNB1* mutations) with underexpression of cell adhesion proteins Putative β-catenin target gene expression > G5 Overexpression of β-catenin > G5 Overexpressed genes: *GPR49, GLUL, PAP/HIP*, LEF1 Underexpressed genes: *CDH1*, highly correlated with downregulation of E-cadherin protein

Figure 14. G1-G6 groups of HCC and its application to the cohort clustering according to the LFQ intensities of 87 above mentioned proteins of samples 01-95 (except 17). Mutations found are indicated with the name of the mutated gene in the lower part of the figure (detailed information can be found in Table 2). Different colours represent different groups of HCC, according to the classification of Boyault (Boyault, et al., 2007). Samples which are not labeled with any colour did not fit into any of the suggested groups. Samples labeled with two colours fitted into both groups of HCC. Groups G1 and G4 are not labeled with colours because none of samples from the cohort fitted into G1 or G4.

Classification based on proteome

In the current study I suggest an alternative grouping according to CCM proteins' expression which is presented in Figure 15**.** It shows the separation of different liver disorders based on the proteome data with the focus on 87 CCM enzymes listed above. Group A is labeled with a violet colour, Group B is labeled with a red colour. The mutation status of each patient is indicated with the name of mutated gene (detailed information can be found in Table 2). Thus, **Group A** includes proteome data from 16 patients: 32, 70, 69, 71, 31, 27, 28, 30, 88, 82, 19, 20, 77, 85, 25, 86; and **Group B** from 31 patients: 24, 91, 68, 84, 94, 89, 90, 26, 33, 73, 23, 78, 93, 66, 74, 75, 22, 92, 21, 29, 34, 95, 67, 87, 83, 79, 80, 35, 76, 81, 72.

Further detailed proteome and metabolome analysis of CCM required a careful choice of control group. Healthy liver biopsies samples were not available due to ethical reasons. The next section (Control group) will describe the approach I used to define the control group of this study.

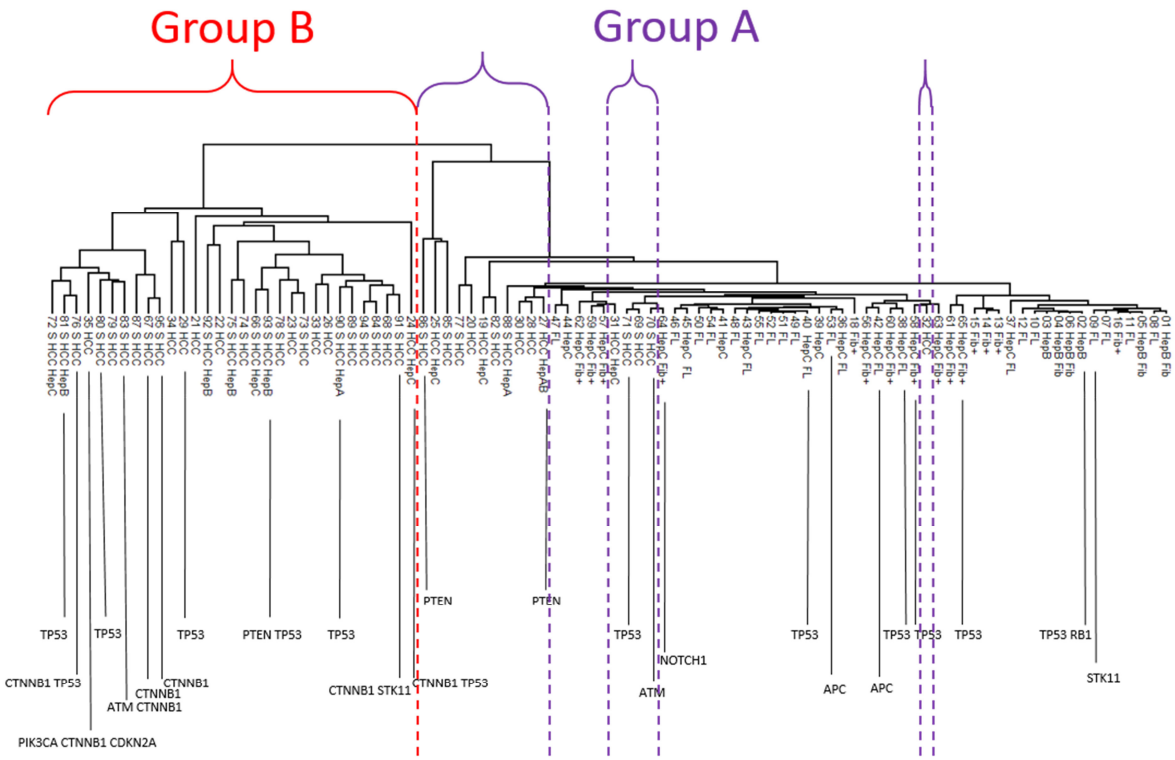


Figure 15. Groups A and B of HCC samples. The separation of different patients based on the proteomic data (LFQ intensities) with the focus on CCM enzymes and suggested grouping according to the proteome clustering. Group A is labeled with a violet, Group B is labeled with a red. The mutation status of each patient is indicated with the name of mutated gene.

Control group

The aim to evaluate specific proteomic and metabolomic changes in CCM of HCC required a relevant control group. Healthy biopsy controls are difficult to obtain due to several ethical issues. Initially our experimental design included fatty liver (FL) samples as a control group, which, in our opinion, was the least severe condition in cohort. The group of FL samples is rather heterogeneous by gender and age of patients, presence of SNPs, and associated diseases. Therefore we performed an additional data analysis to evaluate the appropriate samples for the control group and confirm that FL group is indeed a valid control.

Table 3. Possible control groups from the human cohort comprised of fatty liver samples. Table shows possible control groups for the cohort: the group of fatty liver samples with no mutations and SNPs found consisted of 3 samples (dark green), which is not a sufficient amount for a valid control group. The other FL samples either did not have enough DNA for sequencing (#47), or contained SNPs and hotspots. The samples with confirmed potentially carcinogenic mutations (#9, 53) were excluded from the control group, as well as the sample #47 with no data available. Two different groups of possible controls were compared: [11, 12, 51] 'small' group (dark green), in which no mutations were found in the samples and [7, 8, 10, 11, 12, 46, 48, 49, 50, 51, 52, 54, 55] 'large' group (pale green), which also contains the 'small' group in addition to samples with SNPs in order to validate that the 'large' can be used as a control group.

Sample Code	Diagnosis	Mutations and SNPs	Mutations confirmed	Age	Gender
7	Fatty Liver	TP53 PIK3CA KDR		47	f
8	Fatty Liver	TP53 MET		49	m
9	Fatty Liver	STK11 KDR	STK11	59	f
10	Fatty Liver	PIK3CA KDR		49	f
11	Fatty Liver			43	m
12	Fatty Liver			40	m
46	Fatty Liver, Fibrosis	TP53		46	m
47	Fatty Liver, Fibrosis	na		30	m
48	Fatty Liver, Fibrosis	SMARCB KDR		48	m
49	Fatty Liver	TP53		51	m
50	Fatty Liver	TP53 PIK3CA KDR		31	m
51	Fatty Liver			64	f
52	Fatty Liver, Fibrosis	TP53 KDR		58	f
53	Fatty Liver	TP53 APC	APC	55	m
54	Fatty Liver, Fibrosis	KDR MET		46	m
55	Fatty Liver	TP53 KDR		36	m

Figure 16 shows the correlation of the values taken from multiple CCM enzymes (60 proteins out of 91; proteins which were not detected in one of the groups were excluded from this analysis), the x axes contains the FL small group/HCC values, the y axes contains the FL large group/HCC values; y=0.92x+0.03, R2=0.96 of the trend line, confirming the high homogeneity of both groups.

Based on the performed analysis the control group for further CCM proteome analysis contains samples: [7, 8, 10, 11, 12, 46, 48, 49, 50, 51, 52, 54, 55]; this allowed the use of a more heterogeneous and large cohort of control samples and considered the fibrotic and FL background as well as the possible influence of SNPs.

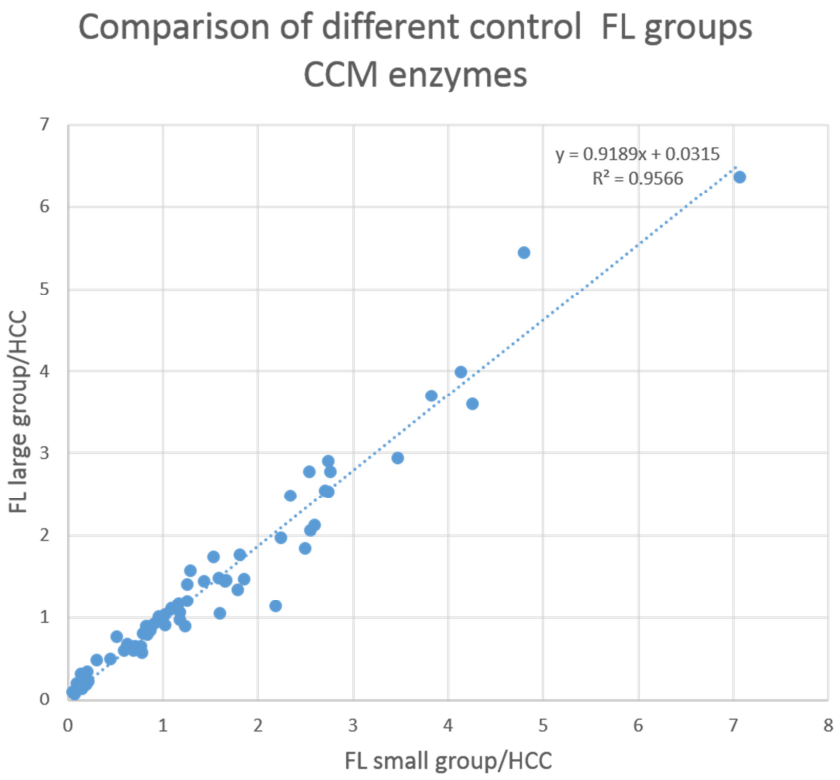


Figure 16. The correlation of CCM proteins (LFQ intensities) between two possible control groups. The correlation of the values taken from multiple CCM enzymes (60 proteins out of 91, excluding proteins which were not detected in one of the groups), the x axes contains the FL small group/HCC values, the y axes contains the FL large group/HCC values; y=0.92x+0.03, R^2=0.96 of the trend line.

FL small group – a group of Fatty Liver samples where no mutations or SNPs were found. The group includes samples 11, 12, 51.

FL large group – a group of Fatty Liver samples with detected SNPs or with no mutations detected. The group includes samples 7, 8, 10, 11, 12, 46, 48, 49, 50, 51, 52, 54, 55.

Common changes in Central Carbon Metabolism of all HCC samples

In order to visualize the proteome and metabolome changes I used Vanted (v. 2.2.1). I used the previously described CCM pathway with the 91 selected key enzymes. Figure 17 represents the colour-coded relative proteomic and metabolomic changes of CCM of the group including all HCC samples compared to the defined control group. At the proteomic level several key features of altered HCC liver metabolism could be underlined.

Glycogen metabolism was downregulated: glycogen synthase, liver form (GYS2) was 2.5 times lower in HCC (t-test, p=0.008), glycogen phosphorylase, liver form (PYGL) was 2.8 times lower in HCC (t-test, $p=4.6*10^{-7}$) (Figure 18).

Glucose *de novo* synthesis was downregulated: fructose-1,6-bisphosphatase 1 (FBP1) was 2.9 times lower in HCC (t-test, $p=1.13*10^{-7}$) (Figure 18).

Transketolase (TKT) enzyme, involved in pentose phosphate pathway was 4.5 higher in HCC samples (t-test, $p=1.09*10^{-3}$) (Figure 18). Several enzymes of glycolysis are upregulated, for example hexokinase 1 (HK1) was not detected in any of control samples, but was present in HCC samples (Figure 18). Glycolysis also showed an isoform switch at the level of phosphoglycerate mutases, aldolases, enolases, lactate dehydrogenases, and pyruvate kinases. For example, Figure 19 shows that aldolase A (ALDOA) was expressed 7.4 times higher in HCC (t-test, $p=5.12*10^{-4}$), and aldolase B (ALDOB) 3 times lower in HCC (t-test, $p=4.49*10^{-8}$); enolase 3 (ENO3) was 6.4 times higher in control samples (t-test, $p=2.83*10^{-8}$), D-lactate dehydrogenase (LDHD) was 2.8 fold higher in HCC (t-test, $p=1.71*10^{-6}$).

Figure 20 shows the downregulation of the TCA cycle and glutaminolysis enzymes expression. Cytoplasmic aconitate hydratase (ACO1) was expressed 1.8 times lower in HCC samples (t-test, $p=4.73*10^{-4}$), Succinate dehydrogenase flavoprotein subunit, mitochondrial (SDHA) was 2 times lower in HCC samples (t-test, $p=1.42*10^{-7}$), succinate dehydrogenase iron-sulfur subunit, mitochondrial (SDHB) was 1.5 times lower in HCC (t-test, $p=2.81*10^{-2}$), GPT was 4 times lower in HCC (t-test, $p=1.12*10^{-9}$). An isoform switch was revealed at the level of isocitrate dehydrogenases: isocitrate dehydrogenase, cytoplasmic (IDH1) was expressed at higher intensity in the control 1.4 (t-test, $p=1.21*10^{-2}$), isocitrate dehydrogenase, mitochondrial (IDH2) was not changed significantly, and isocitrate dehydrogenase subunit alpha, mitochondrial (IDH3A) 4.9 times higher in the HCC (t-test, $p=1.68*10^{-3}$), while subunit beta IDH3B and subunit gamma IDH3G were only detected in the HCC samples (Figure 21).

Changes in metabolic picture of CCM in all HCC samples compared to control group are represented in boxplots in Figure 22. The intensities of several metabolites are presented: lactic acid level was 4.5 fold lower in HCC[1] (t-test, p=0.001), citric acid was 3.8 fold higher in HCC (t-test, p=0.01) and succinic acid 1.6 was fold lower in HCC (t-test, p=0.03).

[1] The observation of lactic acid level decreased in HCC contradicts the Warburg effect. The possible reason for this phenomenon is reviewed in the Discussion section.

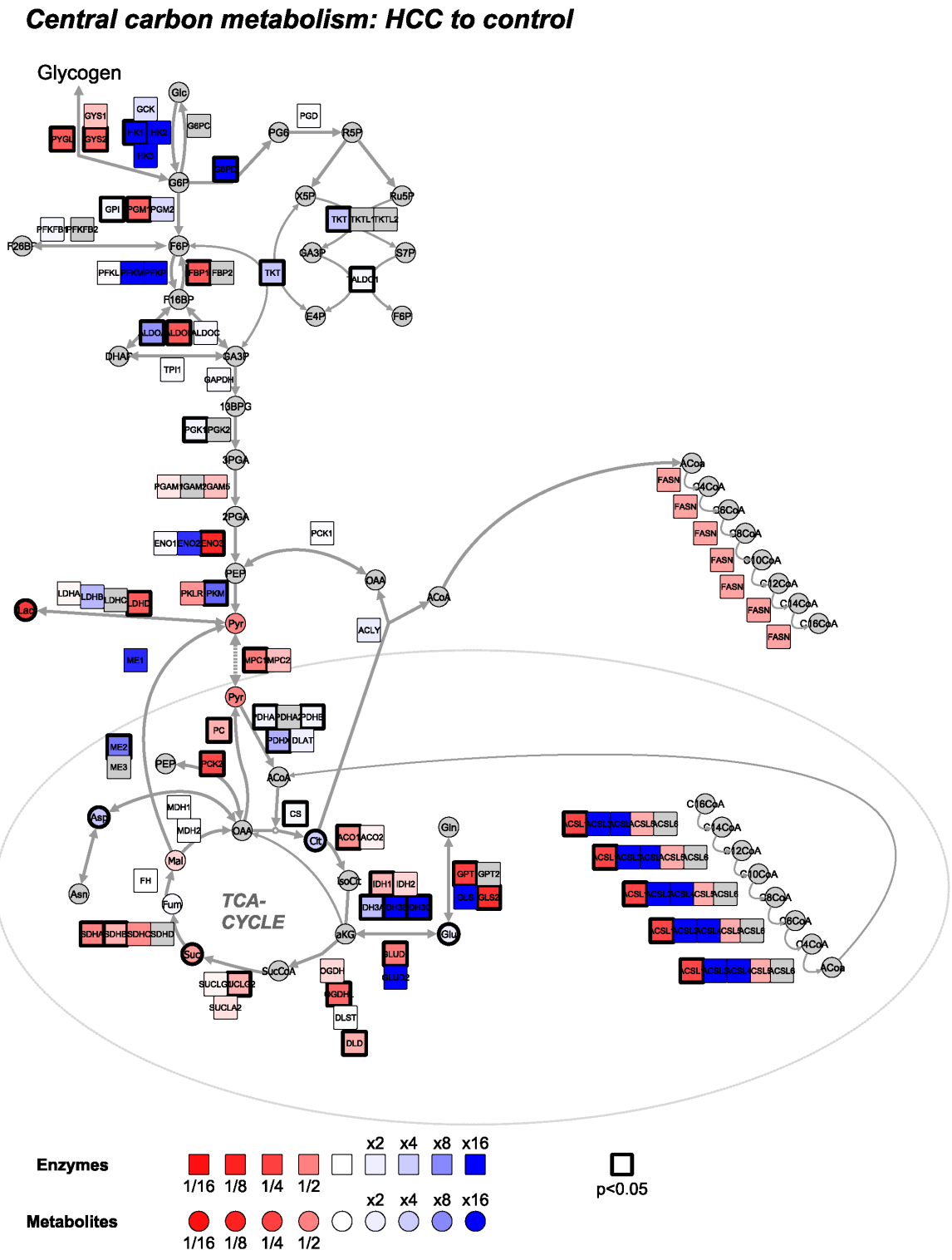


Figure 17. Metabolic changes in all human HCC samples compared to a control group. The colour-coded proteome and metabolome changes in CCM of HCC group compared to a defined control group (FL samples). At the level of proteins glycogen metabolism was downregulated via

GYS2 and PYGL enzymes. GYS2 was 2.5 times lower in HCC group of samples compared to a control group (t-test, p=0.008), PYGL was 2.8 times lower in HCC samples compared to a control group (t-test, p=4.6*10-7). An LFQ intensity of FBP1, an enzyme responsible for glucose *de novo* synthesis, was 2.9 times lower in HCC compared to a control group (t-test, p=1.13*10-7). Changes in pentose-phosphate pathway were characterized by a higher intensity of TKT: it was 4.5 higher in HCC samples compared to a control group (t-test, p=1.09*10-3). A striking feature of glycolytic process was that HK1 was not detected in any of control samples, but was present in HCC samples. Other glycolytic enzymes expressed an isoform switches: phosphoglycerate mutases, aldolases, enolases, lactate dehydrogenases, and pyruvate kinases – these enzymes showed an isoform switch in HCC compared to a control group of samples. Aldolase A was prevalent over aldolase B in HCC samples: ALDOA was expressed 7.4 times higher in HCC compared to a control group of samples (t-test, p=5.12*10-4), and ALDOB was 3 times lower in HCC compared to a control group of samples (t-test, p=4.49*10-8). Enolase 3 was prevalent over enolase 2 in HCC samples: ENO3 was 6.4 times higher in control samples (t-test, p=2.83*10-8). LDHD was prevalent over other lactate dehydrogenase in HCC samples: it's intensity was 2.8 fold higher in HCC compared to a control group (t-test, p=1.71*10-6). TCA cycle enzymes mostly showed the trend of downregulation in HCC samples: an intensity of ACO1 was 1.8 times lower in HCC samples (t-test, p=4.73*10-4), an intensity of SDHA 2 was times lower in HCC samples (t-test, p=1.42*10-7), an intensity of SDHB was 1.5 times lower in HCC compared to a control group of samples (t-test, p=2.81*10-2). Glutamine uptake changes were characterized by GPT enzyme intensity being 4 times lower in HCC compared to control (t-test, p=1.12*10-9). An isoform switch was revealed at the level of isocitrate dehydrogenase subunits: IDH1 intensity was lower in HCC in 1.4 times compared to control (t-test, p=1.21*10-2), and IDH3A intensity was 4.9 times higher in the HCC (t-test, p=1.68*10-3), while IDH3B and IDH3G were only detected in the HCC and not in the control samples. Changes in metabolome were characterized by lactic acid intensity being 4.5 fold lower in HCC compared to control (t-test, p=0.001), citric acid intensity was 3.8 fold higher in HCC compared to control (t-test, p=0.01), succinic acid intensity was 1.6 fold lower in HCC compared to control (t-test, p=0.03). The metabolites are visualized with circles, the proteins are visualized with squares, bold frame indicates the changes with $p<0.05$ (t-test), grey colour of the metabolite (circle) or protein (square) at the scheme indicates that data regarding certain enzyme or metabolite is not available. Red colour indicates decrease of intensity in HCC samples, blue colour indicates increase of intensity in HCC samples.

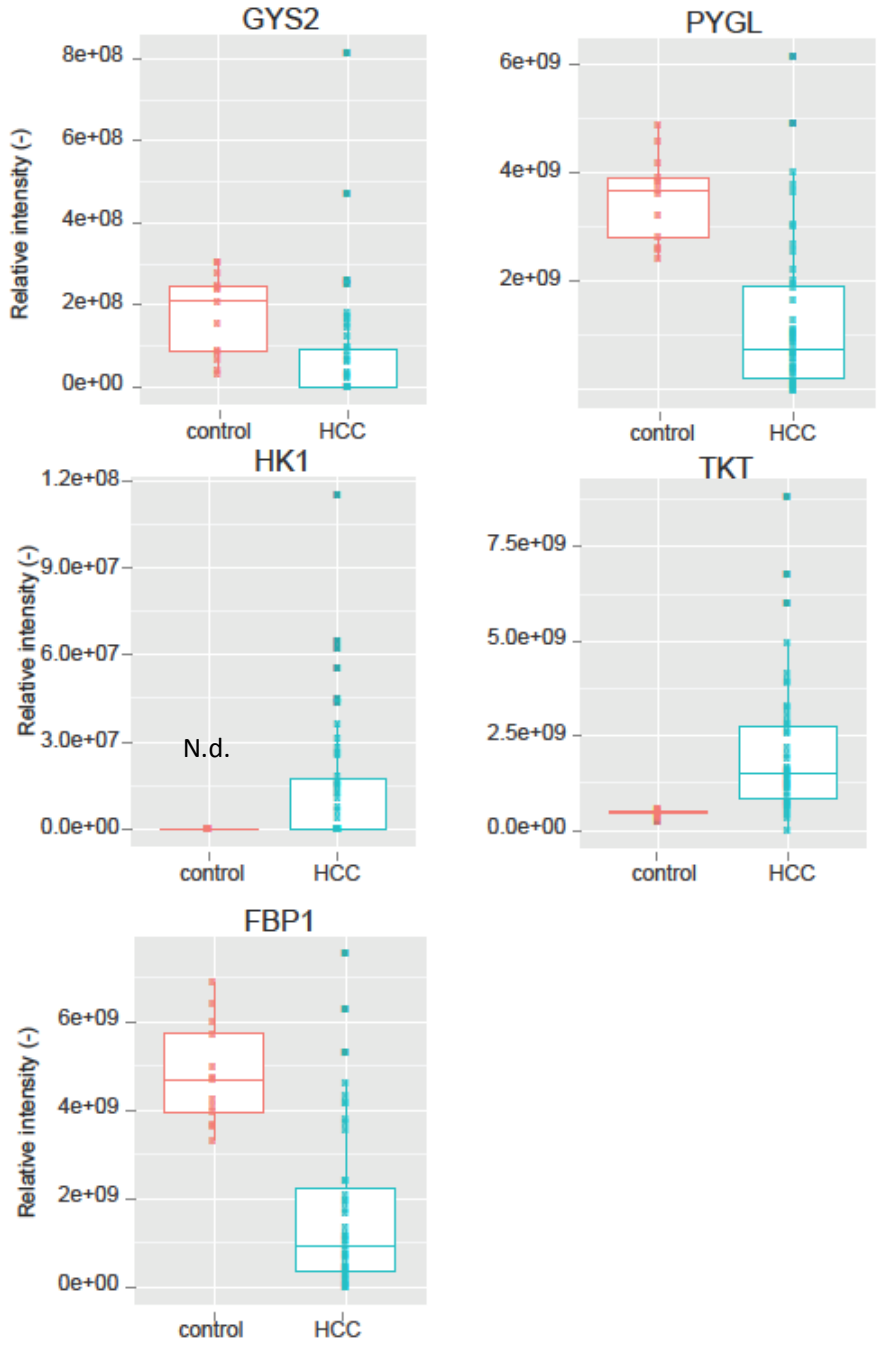


Figure 18. Box plots with the relative intensities of enzymes of glycolysis, gluconeogenesis, glycogen metabolism, and pentose phosphate pathway in all human HCC samples compared to a control group. Glycogen metabolism was downregulated: GYS2 is expressed 2.5 times lower in

HCC (t-test, p=0.008), PYGL was 2.8 times lower in HCC (t-test, $p=4.6*10^{-7}$). Glucose *de novo* synthesis was downregulated: FBP1 was expressed 2.9 times lower in HCC (t-test, $p=1.13*10^{-7}$). TKT enzyme was expressed 4.5 higher in HCC samples (t-test, $p=1.09*10^{-3}$). The enzymes of glycolysis were upregulated, for example HK1 was not detected at any of control samples, but was present at HCC samples. N.d. – not detected.

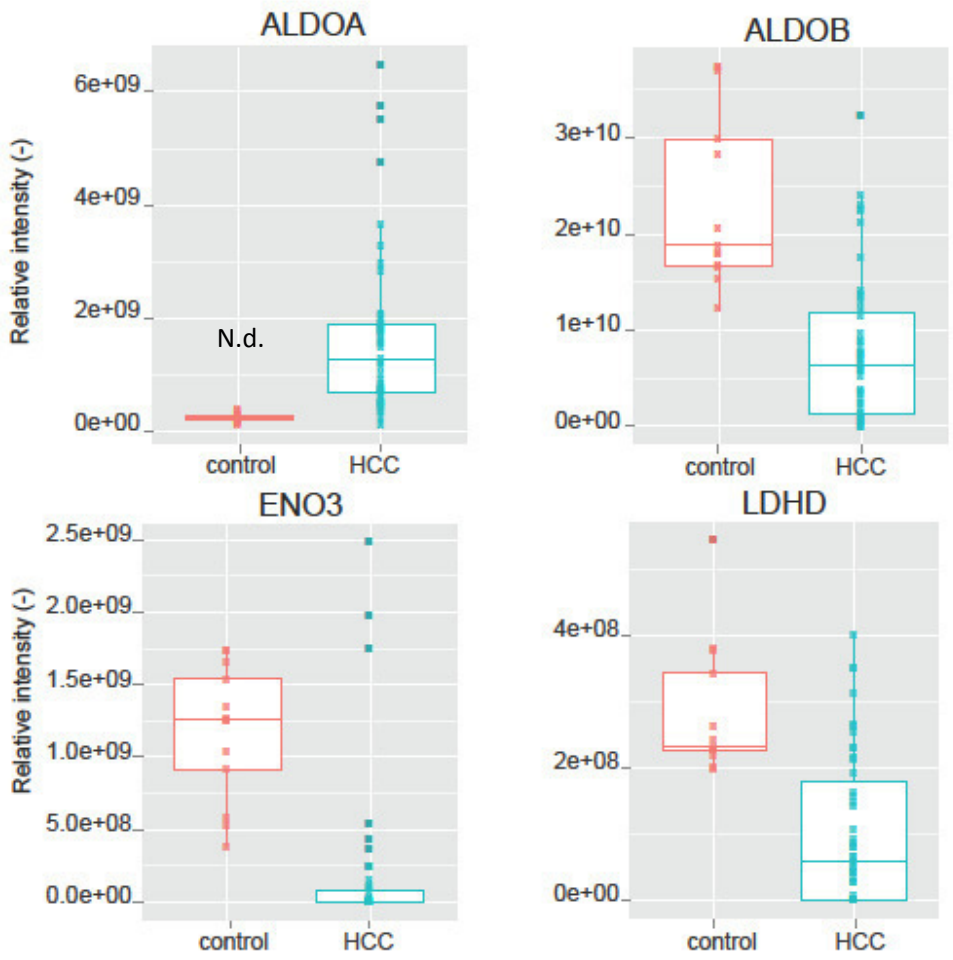


Figure 19. Box plots with the relative intensities of glycolytic enzymes with the detected isoform switch in all human HCC samples compared to a control group. ALDOA was expressed 7.4 times more in HCC (t-test, $p=5.12*10^{-4}$), and ALDOB was 3 times lower in HCC (t-test, $p=4.49*10^{-8}$); ENO3 was 6.4 times higher in control samples (t-test, $p=2.83*10^{-8}$), LDHD was 2.8 fold lower in HCC (t-test, $p=1.71*10^{-6}$). N.d. – not detected.

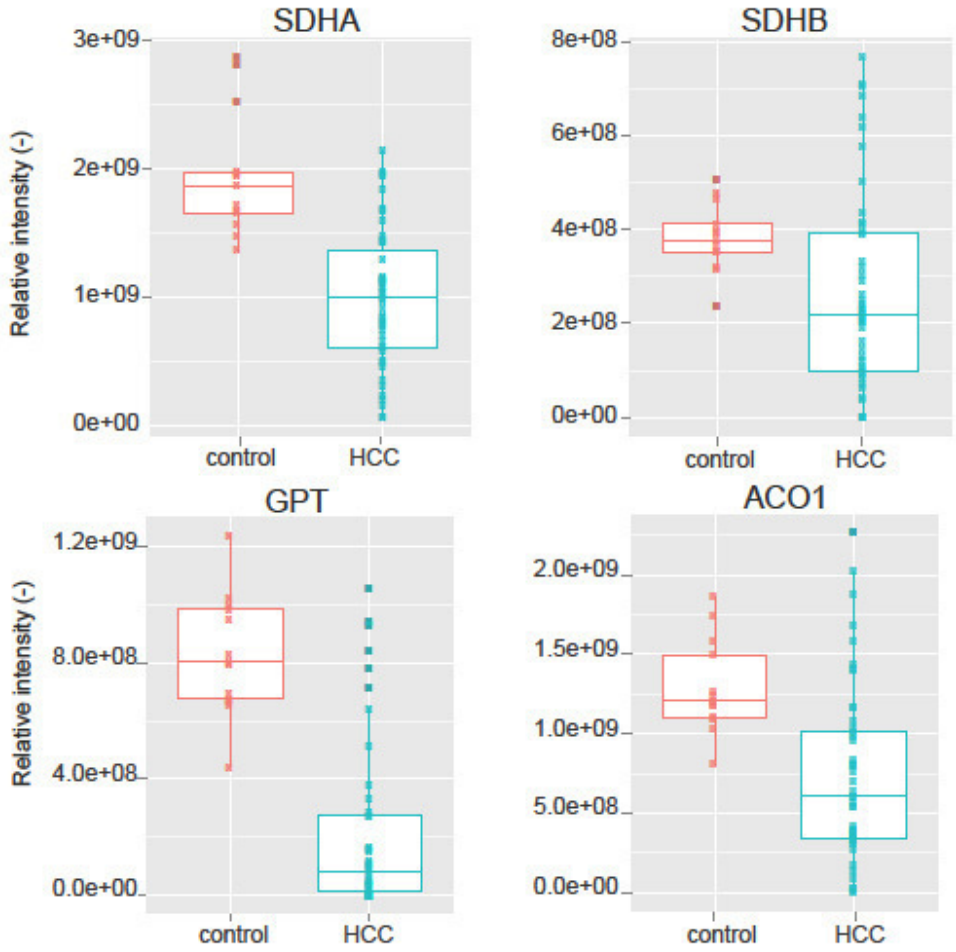


Figure 20. Box plots with the relative intensities of TCA cycle enzymes and enzymes of glutaminolysis in all human HCC samples compared to a control group. ACO1 was expressed 1.8 times lower in HCC samples (t-test, $p=4.7*10^{-4}$), SDHA was 2 times lower in HCC samples (t-test, $p=1.42*10^{-7}$), SDHB was 1.5 times lower in HCC (t-test, $p=2.81*10^{-2}$), GPT was 4 times lower in HCC (t-test, $p=1.12*10^{-9}$).

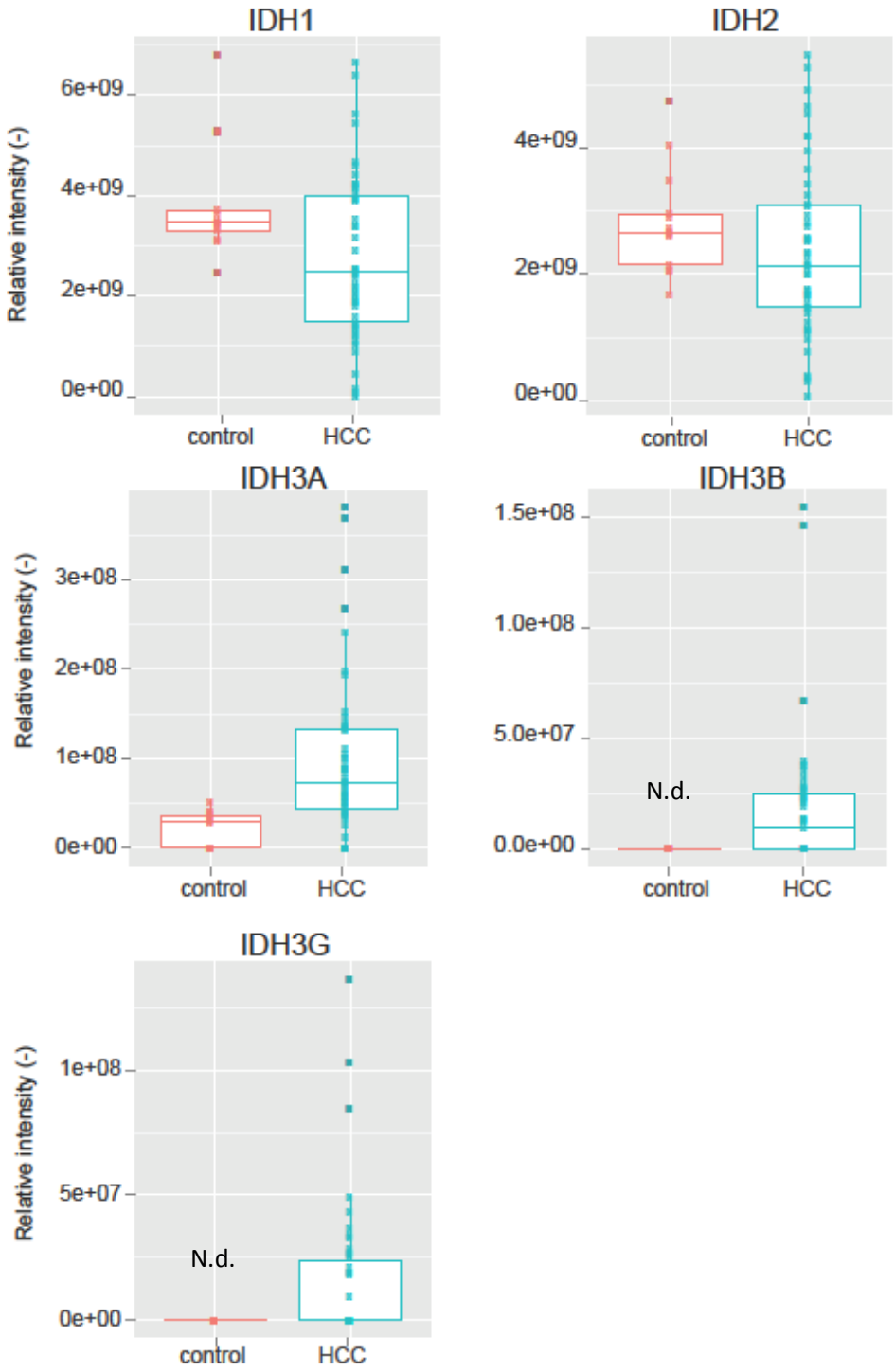


Figure 21. Box plots with the relative intensities of TCA cycle enzymes with revealed isoform switch in all human HCC samples compared to a control group. IDH1 was expressed at higher

intensity in the control 1.4 (t-test, $p=1.21*10^{-2}$), IDH2 was not changed significantly. IDH3A was expressed 4.9 times higher in the HCC (t-test, $p=1.68*10^{-3}$), while IDH3B and IDH3G were only detected in the HCC samples. N.d. – not detected.

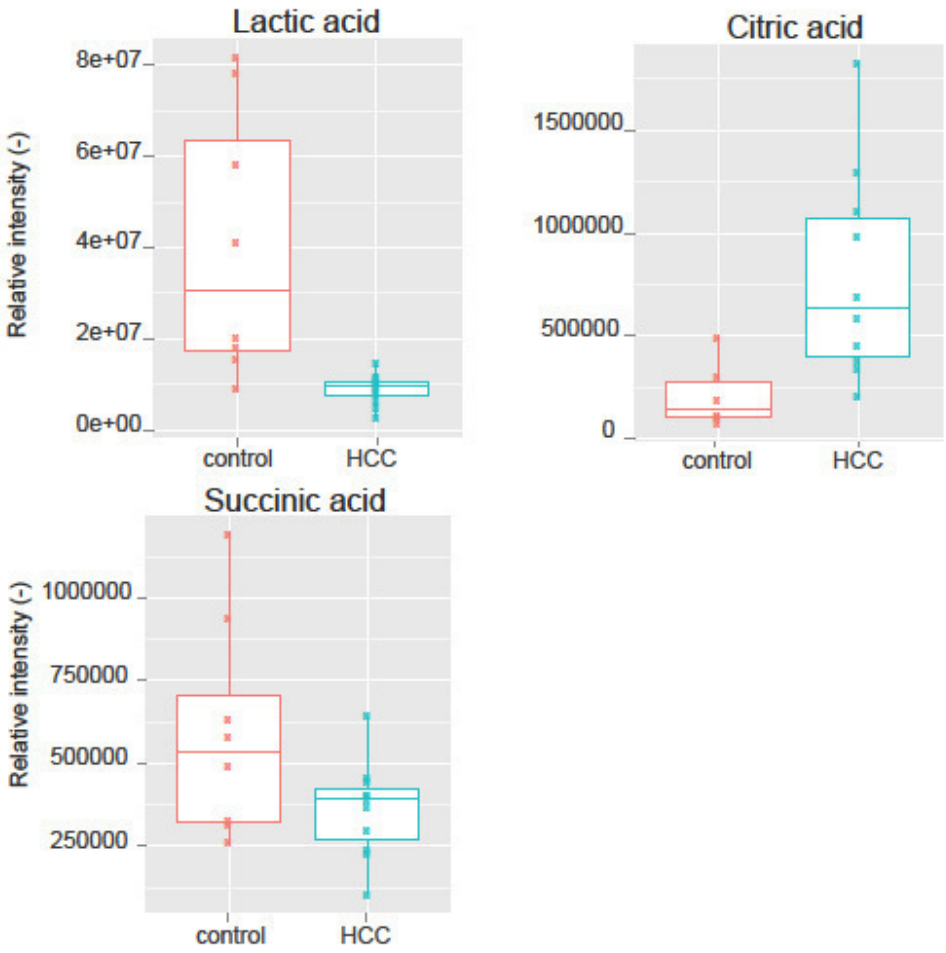


Figure 22. Box plots with the relative intensities of CCM metabolites in all human HCC samples compared to a control group. Lactic acid level was 4.5 fold lower in HCC (t-test, p=0.001), citric acid was 3.8 fold higher in HCC (t-test, p=0.01), succinic acid 1.6 was fold lower in HCC (t-test, p=0.03).

Changes in Central Carbon Metabolism of different HCC groups

The above mentioned focused proteome analysis of samples reveals different subgroups of HCC (Figure 15). I hypothesize that different groups of HCC will show different transformations of CCM. Further I visualize that transformations of CCM (Figure 23 and 29). Group A includes proteome data from 16 patients: 32, 70, 69, 71, 31, 27, 28, 30, 88, 82, 19, 20, 77, 85, 25, 86; and Group B from 31 patients: 24, 91, 68, 84, 94, 89, 90, 26, 33, 73, 23, 78, 93, 66, 74, 75, 22, 92, 21, 29, 34, 95, 67, 87, 83, 79, 80, 35, 76, 81, 72. Only metabolome data obtained from the biopsies was used in the profiles visualization, due to the above mentioned reasons (Figure 13).

HCC Group A

Figure 23 shows the characteristic changes of CCM in human HCC Group A compared to control. At protein expression level glycogen metabolism was downregulated, for example, PYGL was expressed 1.4 times lower in HCC (t-test, p=0.025), while both muscle and liver forms of glycogen synthase (GYS1 and GYS2) were not changed significantly (Figure 24).

Glucose *de novo* synthesis was downregulated: FBP1 was expressed at levels 1.4 times lower in HCC (t-test, p=0.023) (Figure 24).

TKT enzyme was expressed 2.9 fold more in HCC samples (t-test, p=0.012).

The enzymes of glycolysis were upregulated or underwent an isoform switch at the level of phosphoglycerate mutases, aldolases, and lactate dehydrogenases. ALDOA was expressed 3.8 times higher in HCC (t-test, p=0.0001), ALDOB was 1.5 times lower in HCC (t-test, $p<0.05$) (Figure 24), ENO3 was 2.3 times higher in control samples (t-test, p=0.014), pyruvate kinase M (PKM) was 3.8 times higher in HCC samples (t-test, p=0.002) (Figure 25). Lactate dehydrogenases underwent isoform switch: LDHB was expressed 2.5 times more in HCC samples (t-test, p=0.004) and LDHD was 1.5 fold more in the control (t-test, p=0.011) (Figure 25).

Most of the enzymes of the TCA cycle did not show any drastic change in expression: both cytoplasmic and mitochondrial aconitate hydratases (ACO1 and ACO2) (Figure 25) were not changed significantly, as well as malate dehydrogenases, and the majority of isocitrate dehydrogenase subunits (Figure 26).

Glutaminolysis was downregulated at the proteomic level: GPT was expressed 1.8 times less in HCC, and GLS2 was 2.2 times less (t-test, p=0.0014 and p=0.031 respectively) (Figure 27). At the metabolomic level various changes were detected: lactic acid level was 4.2 fold decreased in HCC (t-test, p=0.013), citric acid level was 5 times higher in HCC (t-test, p=0.002), fumaric acid level was

1.6 times higher in HCC (t-test, p=0.024), and glutamic acid level was 2.6 times higher in HCC (t-test, p=0.004) (Figure 28).

Central carbon metabolism: Group A HCC to control

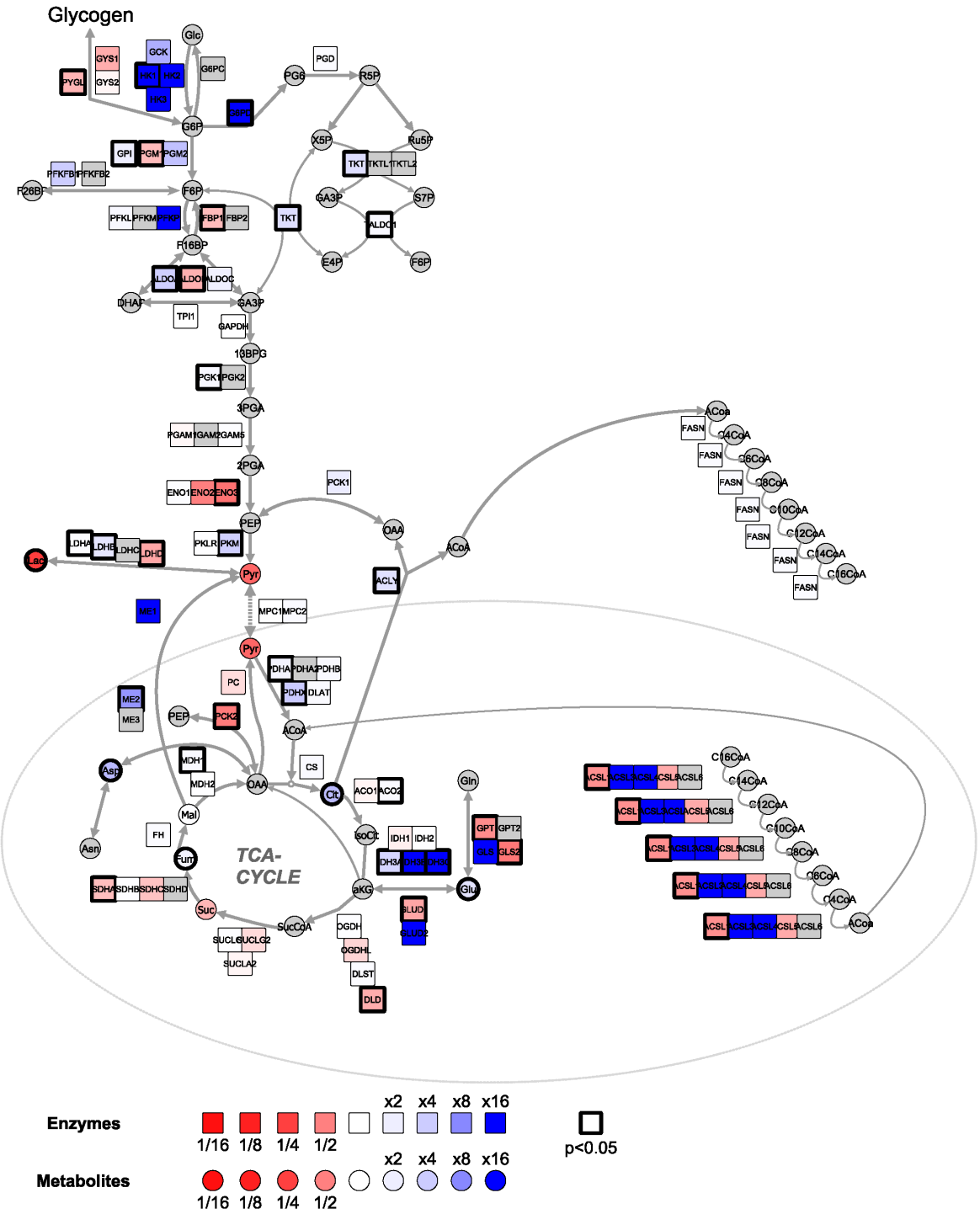


Figure 23. CCM changes in human HCC Group A compared to a control group. The colour-coded relative proteome and metabolome changes in CCM of HCC group A compared to defined control group of FL samples. At the proteomic level glycogen metabolism was downregulated, for example, PYGL was expressed 1.4 times lower in HCC (t-test, p=0.025), while both GYS1 and GYS2 were not changed significantly. Glucose *de novo* synthesis is downregulated: FBP1 was expressed 1.4 times less in HCC (t-test, p=0.023). TKT enzyme was expressed 2.9 fold more in HCC samples (t-test, p=0.012). The enzymes of glycolysis were upregulated or underwent an isoform switch at the level of phosphoglycerate mutases, aldolases, and lactate dehydrogenases. ALDOA was expressed 3.8 times more in HCC (t-test, p=0.0001), and ALDOB 1.5 times less in HCC (t-test, p<0.05), ENO3 2.3 times higher in control samples (t-test, p 0.014), PKM 3.8 times higher in HCC samples (t-test, p=0.002). Lactate dehydrogenases underwent an isoform switch: LDHB was expressed 2.5 higher in HCC samples (t-test, p=0.004) and LDHD 1.5 fold higher in control (t-test, p=0.011). Most of the enzymes of the TCA cycle did not show any drastic change in the expression: both ACO1 and ACO2 were not changed significantly, neither was malate dehydrogenases, or the majority of isocitrate dehydrogenase subunits. Glutaminolysis was downregulated at the proteomic level: GPT was expressed 1.8 times lower in HCC, and GLS2 2.2 times lower (t-test, p=0.0014 and p=0.031 respectively). At the metabolomic level, various changes were detected: the lactic acid level was 4.2 lower in HCC (t-test, p=0.013), the citric acid level was 5 times higher in HCC (t-test, p=0.002), the fumaric acid level was 1.6 times higher in HCC (t-test, p=0.024), and the glutamic acid level was 2.6 times higher in HCC (t-test, p=0.004). The metabolites are visualized with circles, the proteins are visualized with squares, bold frame indicates the changes with p<0.05 (t-test), grey colour of the metabolite (circle) or protein (square) at the scheme indicates that data regarding certain enzyme or metabolite is not available. Red colour indicates decrease of intensity in HCC samples, blue colour indicates increase of intensity in HCC samples.

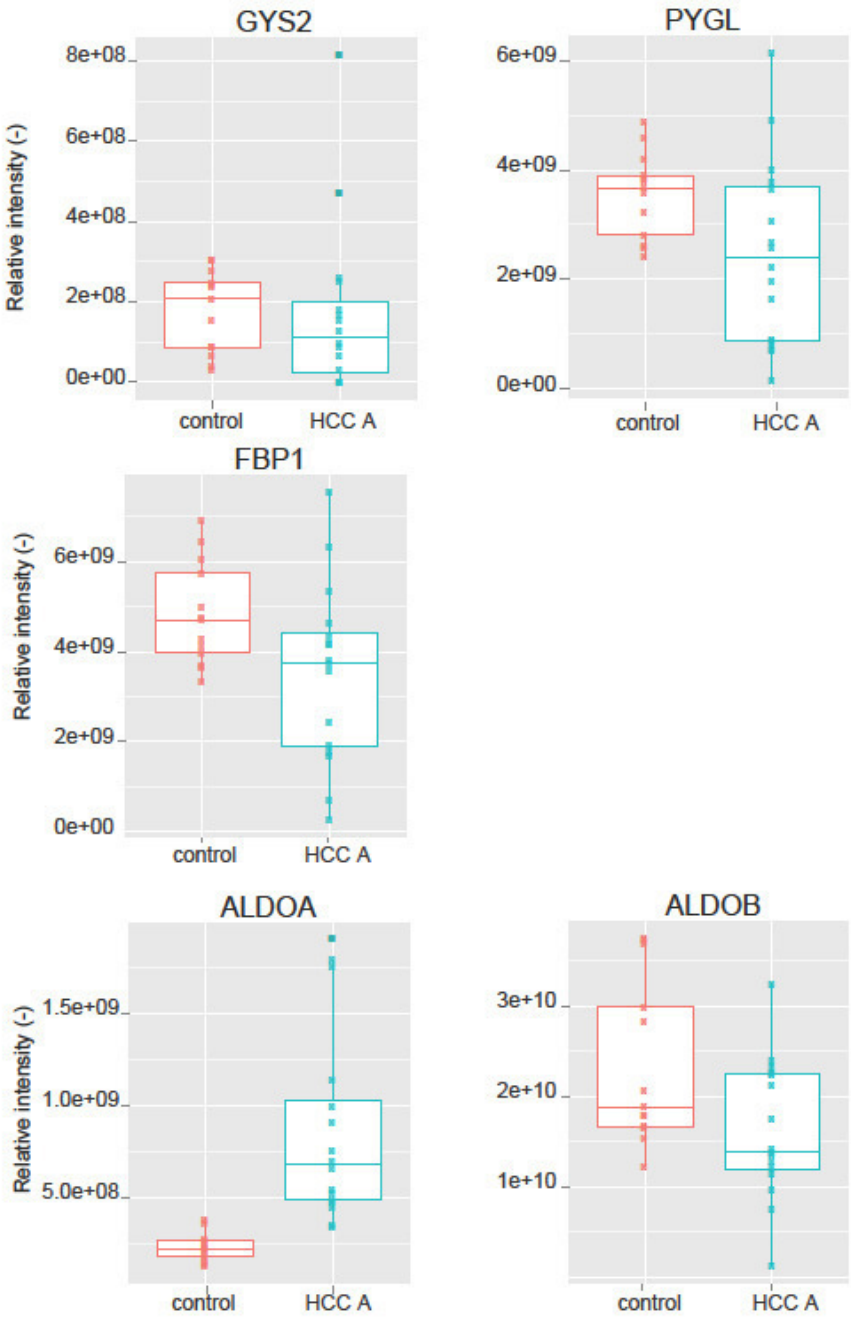


Figure 24. Box plots with the relative intensities of glycolytic and gluconeogenetic enzymes, and enzymes participating in glycogen metabolism in human HCC Group A compared to a control group. PYGL was expressed 1.4 times lower in HCC (t-test, p=0.025), GYS2 was not changed

significantly, FBP1 was expressed 1.4 times less in HCC (t-test, p=0.023), ALDOA was 3.8 times more in HCC (t-test, p=0.0001), and ALDOB was 1.5 times lower in HCC (t-test, p<0.05).

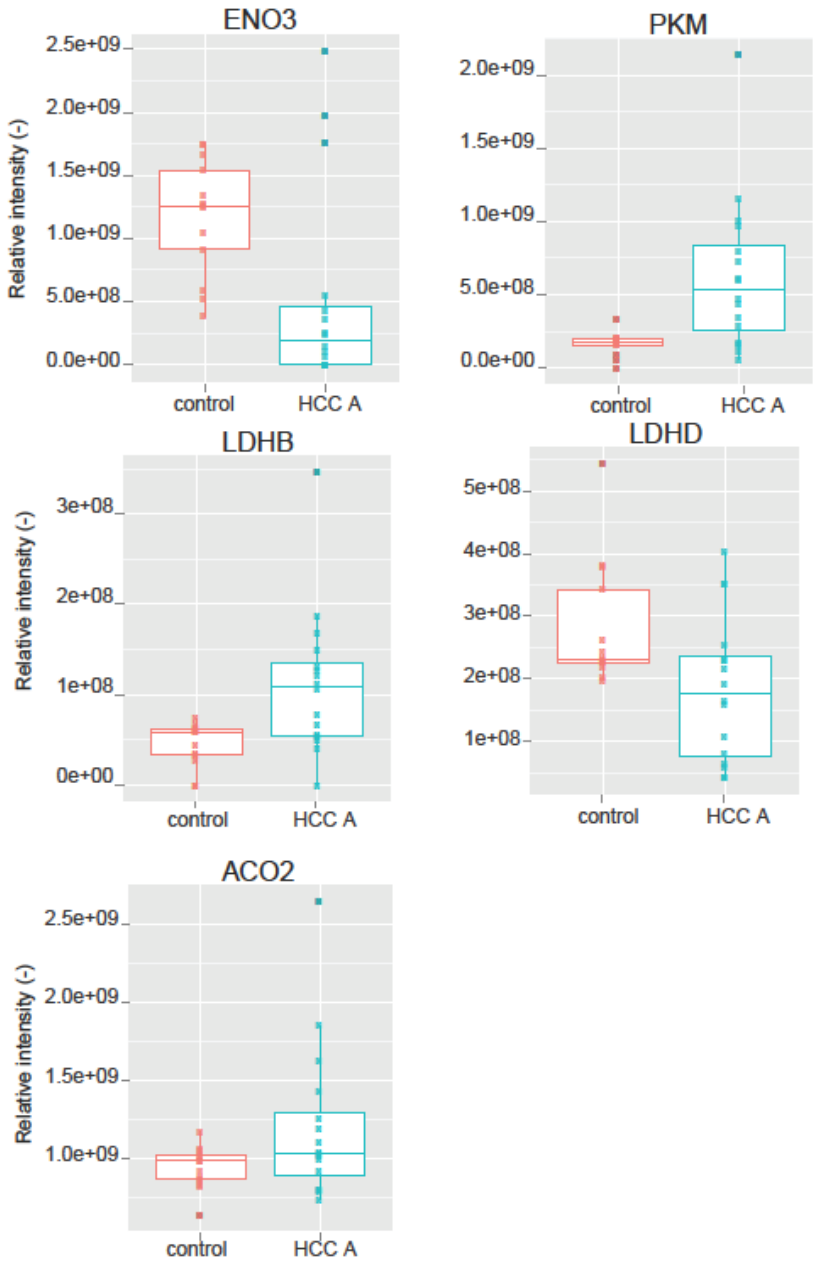


Figure 25. Box plots with the relative intensities of glycolytic and TCA cycle enzymes in human HCC Group A compared to a control group of samples. ENO3 was 2.3 times higher in control samples (t-test, p=0.014), PKM was 3.8 times higher in HCC samples (t-test, p=0.002). Lactate

dehydrogenases undergo isoform switch: LDHB was expressed 2.5 higher in HCC samples (t-test, p=0.004) and LDHD 1.5 fold higher in control (t-test, p=0.011). ACO2 was not changed significantly.

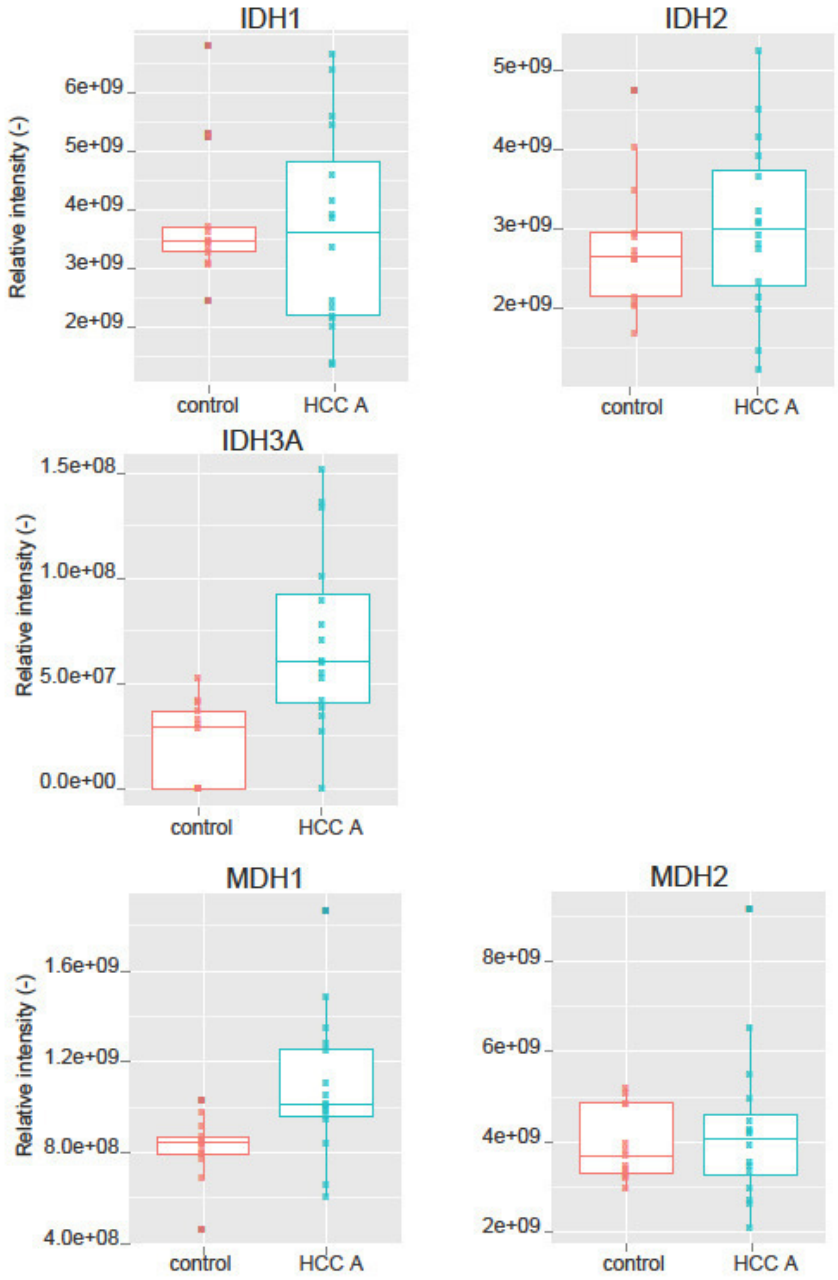


Figure 26. Box plots with the relative intensities of TCA cycle enzymes in human HCC Group A compared to a control group. IDH1 and IDH2 expression was not changed significantly in HCC samples, and IDH3A was expressed significantly higher in HCC (t-test, p=0.0003). MDH1 was

expressed 1.3 fold higher in HCC (t-test, p=0.004) and MDH2 was not significantly changed in the HCC samples.

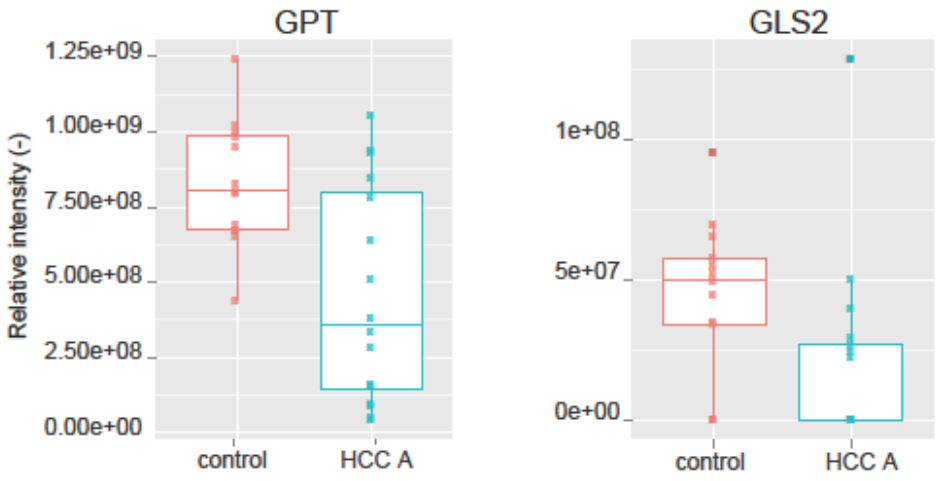


Figure 27. Box plots with the relative intensities of enzymes of glutaminolysis in human HCC Group A compared to a control group. GPT was expressed 1.8 times lower in HCC (t-test, p=0.001) and GLS2 was 2.2 times lower in HCC (t-test, p=0.031).

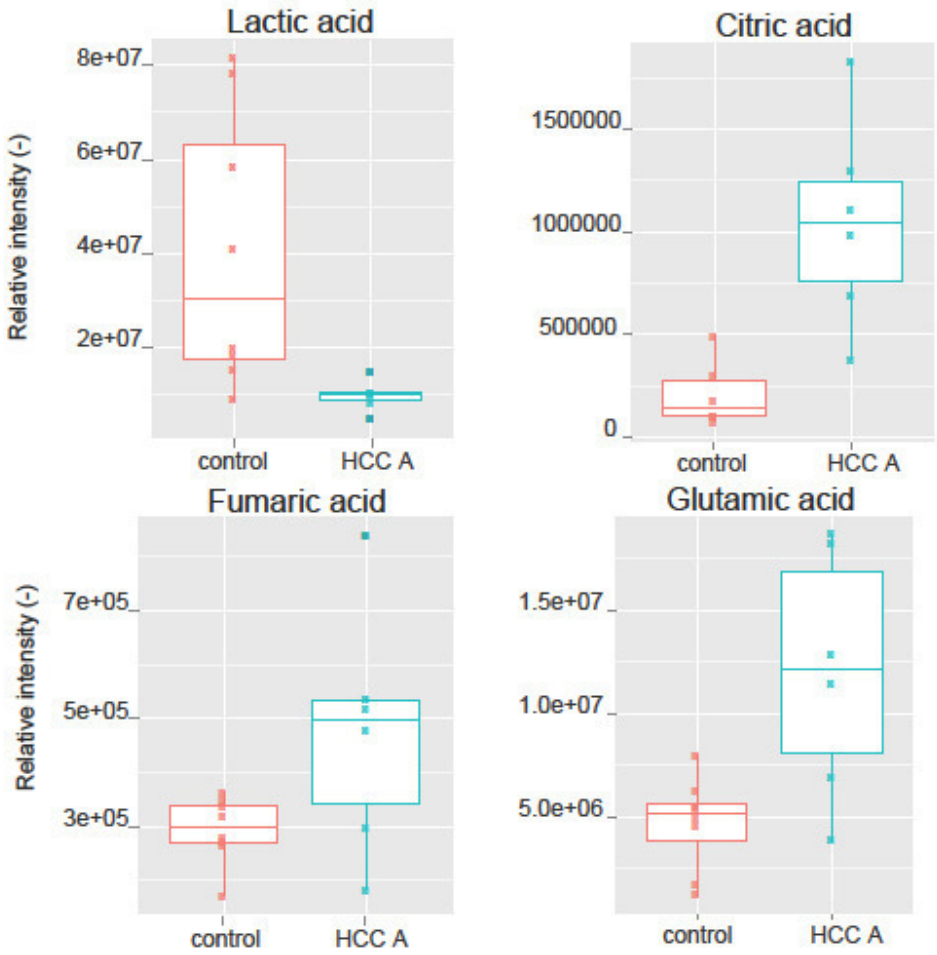


Figure 28. Box plots with the relative intensities of CCM metabolites in in human HCC Group A compared to a control group. The lactic acid level was 4.2 lower in HCC (t-test, p=0.013), citric acid level was 5 times higher in HCC (t-test, p=0.002), fumaric acid level was 1.6 times higher in HCC (t-test, p=0.024), and glutamic acid level is 2.6 times higher in HCC (t-test, p=0.004).

HCC Group B

Figure 29 represents the changes of CCM in the HCC Group B compared to control. At the proteomic level glycogen metabolism was downregulated, for examples, PYGL was expressed 5.4 times lower in HCC (t-test, $p=1.16*10^{-15}$), GYS2 8.7 times lower in HCC (t-test, $p=4.78*10^{-9}$) (Figure 30).

Glucose *de novo* synthesis was downregulated: FBP1 was expressed 6.9 times lower in HCC (t-test, $p=4.46*10^{-19}$) (Figure 30).

TKT was expressed 5.2 times higher in HCC (t-test, $p=3.12*10^{-4}$), and glucose-6-phosphate 1-dehydrogenase (G6PD), which is also involved in pentose phosphate pathway, was not detected in the control samples (Figure 30).

Glycolysis in HCC was characterized by an isoform switch at the levels of phosphoglycerate mutases, aldolases, enolases, pyruvate kinases, and lactate dehydrogenases. ALDOA was expressed 9.2 times higher in HCC (t-test, $p=1.24*10^{-4}$), ALDOB 6.3 times lower in HCC (t-test, $p=5.31*10^{-14}$), ALDOC was not changed (Figure 31). ENO1 was expressed 1.3 fold higher in HCC (t-test, $p=4.84*10^{-2}$), ENO3 was 122 times lower in HCC (t-test, $p=5.43*10^{-18}$) (Figure 31). PKM was expressed 13 times higher in HCC samples (t-test, $p=2.95*10^{-3}$), and PKLR 3 times lower in HCC (t-test, $p=7.14*10^{-4}$) (Figure 32). LDHB was expressed 6.6 times higher in HCC samples (t-test, $p=3.81*10^{-2}$), LDHD was 4.7 fold higher in control (t-test, $p=3.79*10^{-9}$), LDHA was not changed significantly (Figure 32).

65% of the enzymes of the TCA cycle were downregulated: ACO1 was expressed 2.7 and ACO2 1.3 times lower (t-test, $p=6.43*10^{-9}$ and $p=4.31*10^{-2}$ respectively) (Figure 33). IDH showed a switch within subunits: IDH1 was expressed 1.7 times lower in HCC (t-test, $p=4.63*10^{-4}$), IDH3A was 5.6 times higher in HCC (t t-test, $p=1.65*10^{-3}$), IDH3B was only detected in HCC (Figure 33).

Glutaminolysis was downregulated at the proteomic level: GPT was expressed 10.5 times lower in HCC (t-test, $p=1.11*10^{-17)}$, and GLS2 was not detected in HCC samples (with the exception of 1 sample) (Figure 34).

At the metabolomic level various changes were detected: the lactic acid level was 4.9 lower in HCC (t-test, p=0.018), the malic acid level was 2.7 times lower in HCC (t-test, p=0.004), and the succinic acid level was 2.1 times lower in HCC (t-test, p=0.04) (Figure 35).

Central carbon metabolism: Group B HCC to control

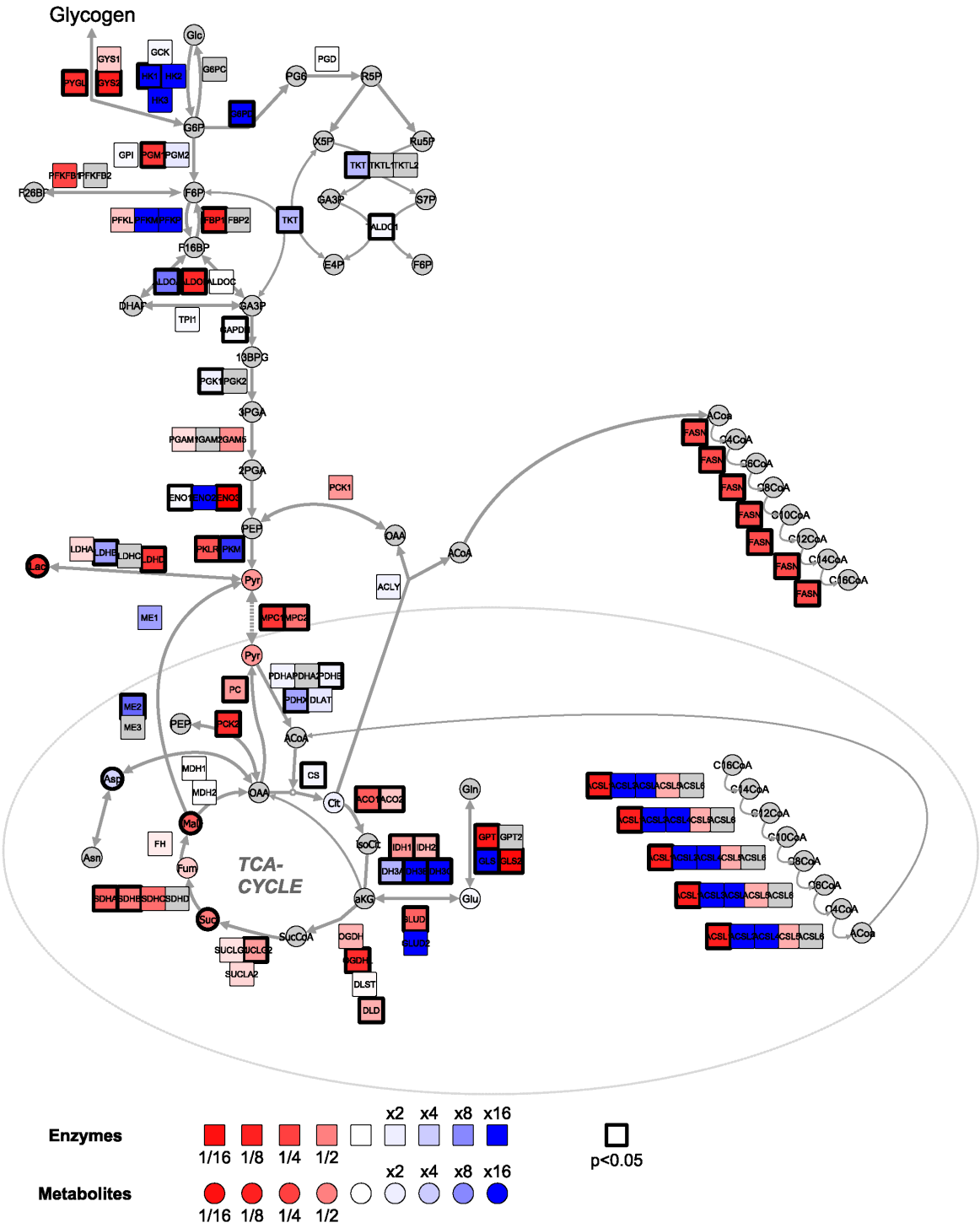


Figure 29. CCM changes in human HCC Group B compared to a control group. At the proteomic level glycogen metabolism was downregulated, for examples, PYGL was expressed 5.4 times lower in HCC (t-test, $p=1.16*10^{-15}$), GYS2 is 8.7 times less in HCC (t-test, $p=4.78*10^{-9}$). Glucose *de novo* synthesis was downregulated: FBP1 is expressed 6.9 times lower in HCC (t-test, $p=4.46*10^{-19}$). TKT

was expressed 5.2 times higher in HCC (t-test, $p=3.12*10^{-4}$), and G6PD, which is also involved in pentose phosphate pathway, was not detected in control samples. Glycolysis was characterized by an isoform switch at the levels of phosphoglycerate mutases, aldolases, enolases, pyruvate kinases, and lactate dehydrogenases. ALDOA was expressed 9.2 times higher in HCC (t-test, $p=1.24*10^{-4}$), ALDOB was 6.3 times less in HCC (t-test, $p=5.31*10^{-14}$), ALDOC was not changed. ENO1 was expressed 1.3 fold higher in HCC (t-test, $p=4.84*10^{-2}$), ENO3 was 122 times lower in HCC (t-test, $p=5.43*10^{-18}$). PKM was expressed 13 times higher in HCC samples (t-test, $p=2.95*10^{-3}$), and PKLR was 3 times lower in HCC (t-test, $p=7.14*10^{-4}$). LDHB was expressed 6.6 times higher in HCC samples (t-test, $p=3.81*10^{-2}$), LDHD was 4.7 fold higher in control (t-test, $p=3.79*10^{-9}$), LDHA was not changed significantly. 65% of the enzymes of the TCA cycle are downregulated: ACO1 was expressed 2.7 and ACO2 was 1.3 times lower (t-test, $p=6.43*10^{-9}$ and $p=4.31*10^{-2}$ respectively). IDH showed a switch between subunits: IDH1 was expressed 1.7 times lower in HCC (t-test, $p=4.63*10^{-4}$), IDH3A 5.6 times higher in HCC (t t-test, $p=1.65*10^{-3}$), IDH3B was only detected in HCC. Glutaminolysis was downregulated at the proteomic level: GPT was expressed 10.5 times lower in HCC (t-test, $p=1.11*10^{-17)}$, and GLS2 was not detected in HCC samples (with the exception of 1 sample). At the metabolomic level various changes were detected: the lactic acid level was 4.9 lower in HCC (t-test, p=0.018), the malic acid level was 2.7 times lower in HCC (t-test, p=0.004), and the succinic acid level was 2.1 times lower in HCC (t-test, p=0.04). The metabolites are visualized with circles, the proteins are visualized with squares, bold frame indicates the changes with $p<0.05$ (t-test), grey colour of the metabolite (circle) or protein (square) at the scheme indicates that data regarding certain enzyme or metabolite is not available. Red colour indicates decrease of intensity in HCC samples, blue colour indicates increase of intensity in HCC samples.

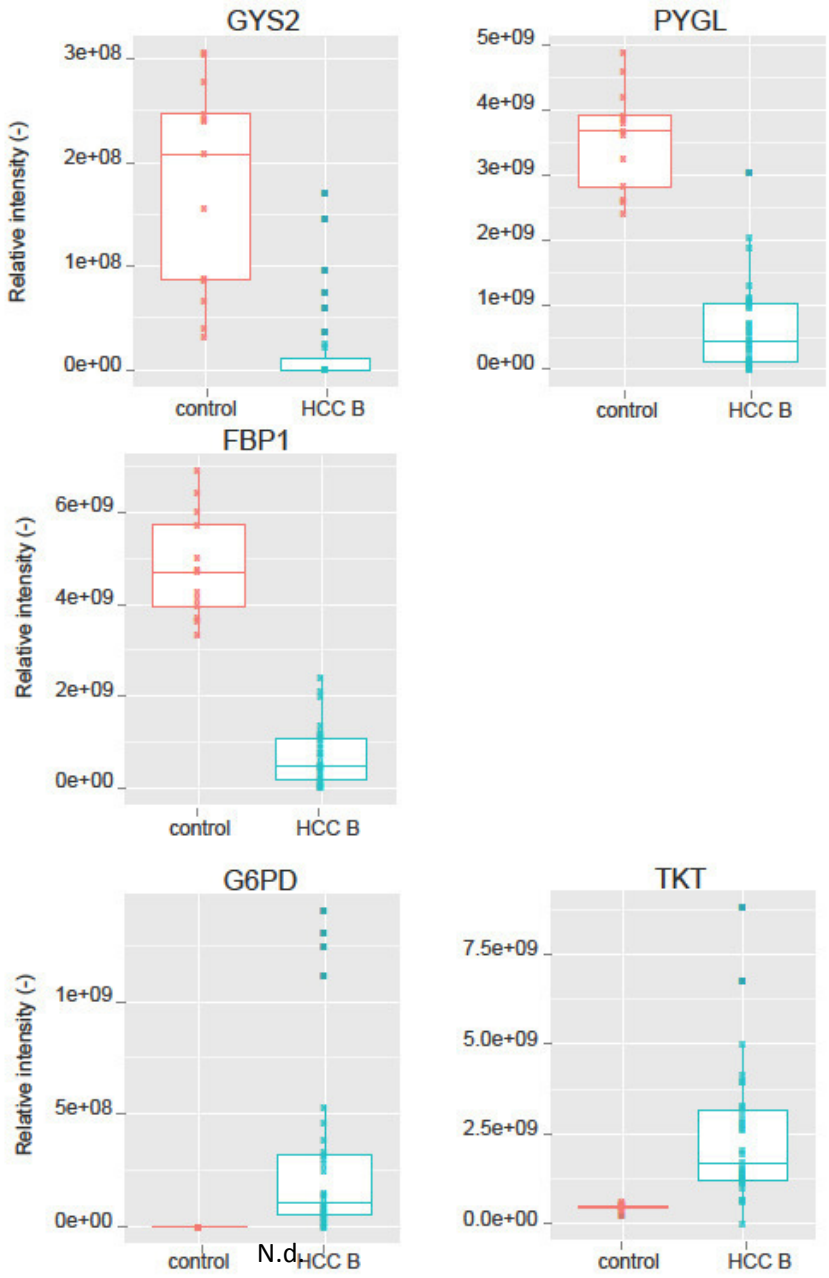


Figure 30.

Box plots with the relative intensities of enzymes of gluconeogenesis, glycogen metabolism, and pentose phosphate pathway in human HCC Group B compared to a control group. PYGL was expressed 5.4 times less in HCC (t-test, $p=1.16*10^{-15}$), GYS2 was 8.7 times more in HCC (t-test, $p=4.78*10^{-9}$). FBP1 was expressed 6.9 times less in HCC (t-test, $p=4.46*10^{-19}$). TKT was expressed

5.2 times higher in HCC (t-test, $p=3.12*10^{-4}$), and G6PD was not detected in control samples. N.d. – not detected.

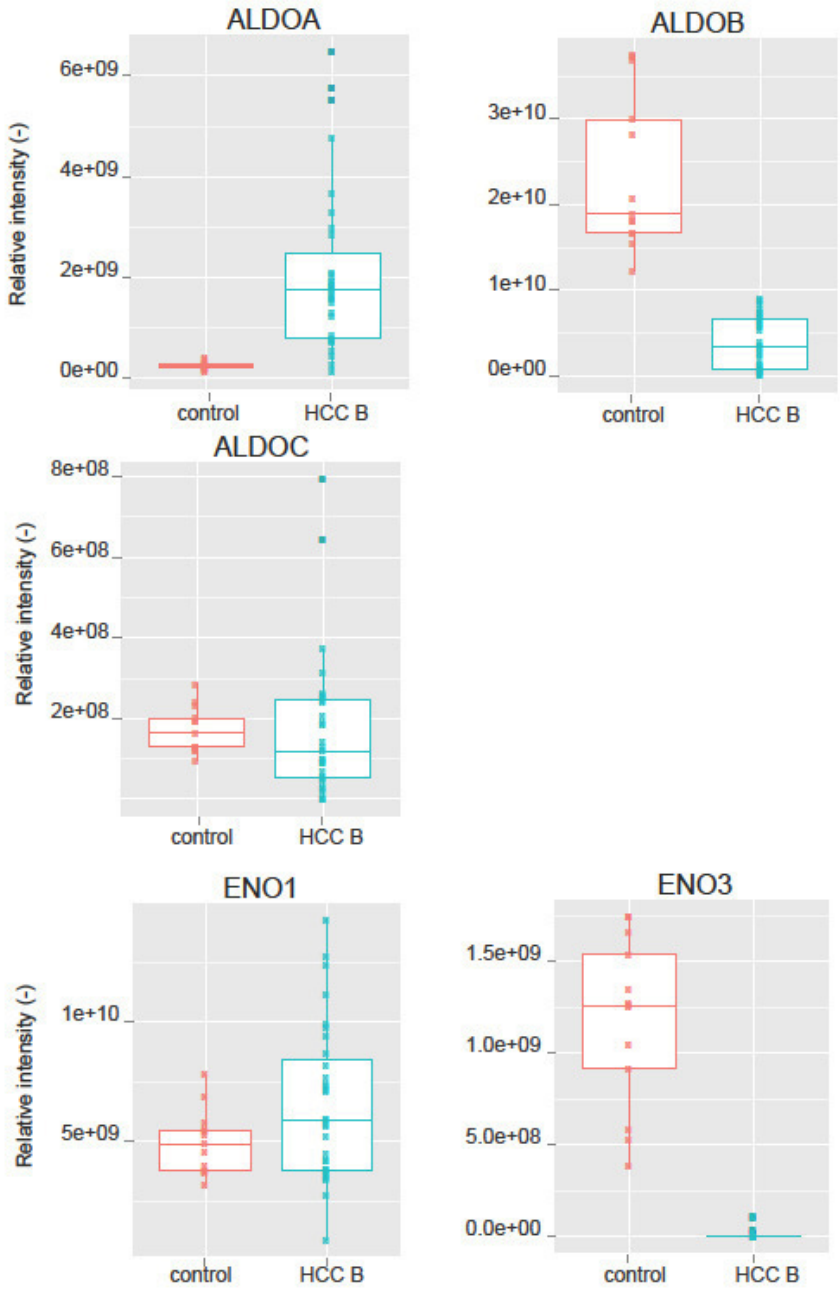


Figure 31. Box plots with the relative intensities of glycolytic enzymes in human HCC Group B compared to a control group (aldolases, enolases). ALDOA was expressed 9.2 times more in HCC

(t-test, $p=1.24*10^{-4}$), ALDOB was 6.3 times less in HCC (t-test, $p=5.31*10^{-14}$), ALDOC was not changed. ENO1 was expressed 1.3 fold more in HCC (t-test, $p=4.84*10^{-2}$), ENO3 was 122 times less in HCC (t-test, $p=5.43*10^{-18}$).

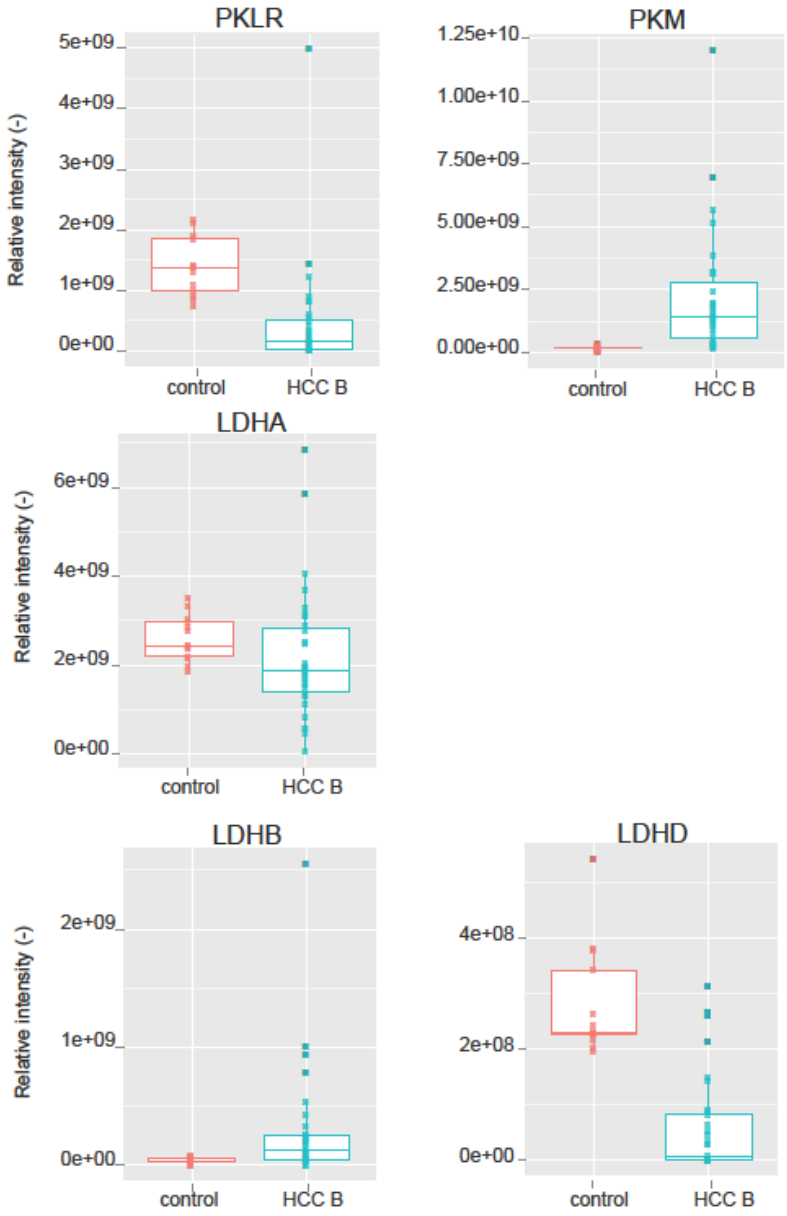


Figure 32. Box plots with the relative intensities of glycolytic enzymes in human HCC Group B compared to a control group (pyruvate kinases, lactate dehydrogenases). PKM was expressed 13 times more in HCC samples (t-test, $p=2.95*10^{-3}$), and PKLR was 3 times less in HCC (t-test,

p=7.14*10^{-4}). LDHB was expressed 6.6 times more in HCC samples (t-test, p=3.81*10^{-2}), and LDHD 4.7 fold more in control (t-test, p=3.79*10^{-9}), LDHA was not changed significantly.

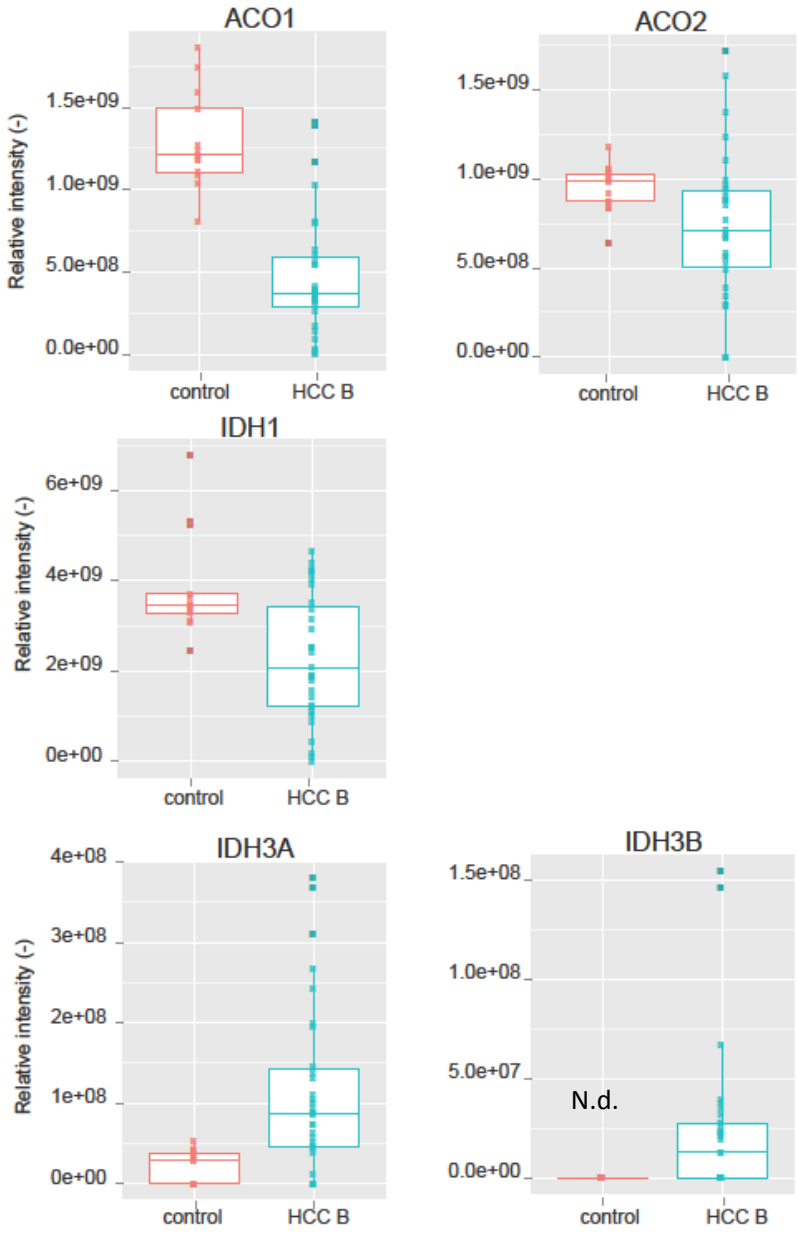


Figure 33. Box plots with the relative intensities of TCA cycle enzymes in human HCC Group B compared to a control group. ACO1 was expressed 2.7 less and ACO2 1.3 times less in HCC (t-test, p=6.43*10^{-9} and p=4.31*10^{-2}, respectively). IDH1 was expressed 1.7 times less in HCC (t-test,

p=4.63*10^{-4}), IDH3A was 5.6 times more in HCC (t t-test, p=1.65*10^{-3}), IDH3B was only detected in HCC. N.d. – not detected.

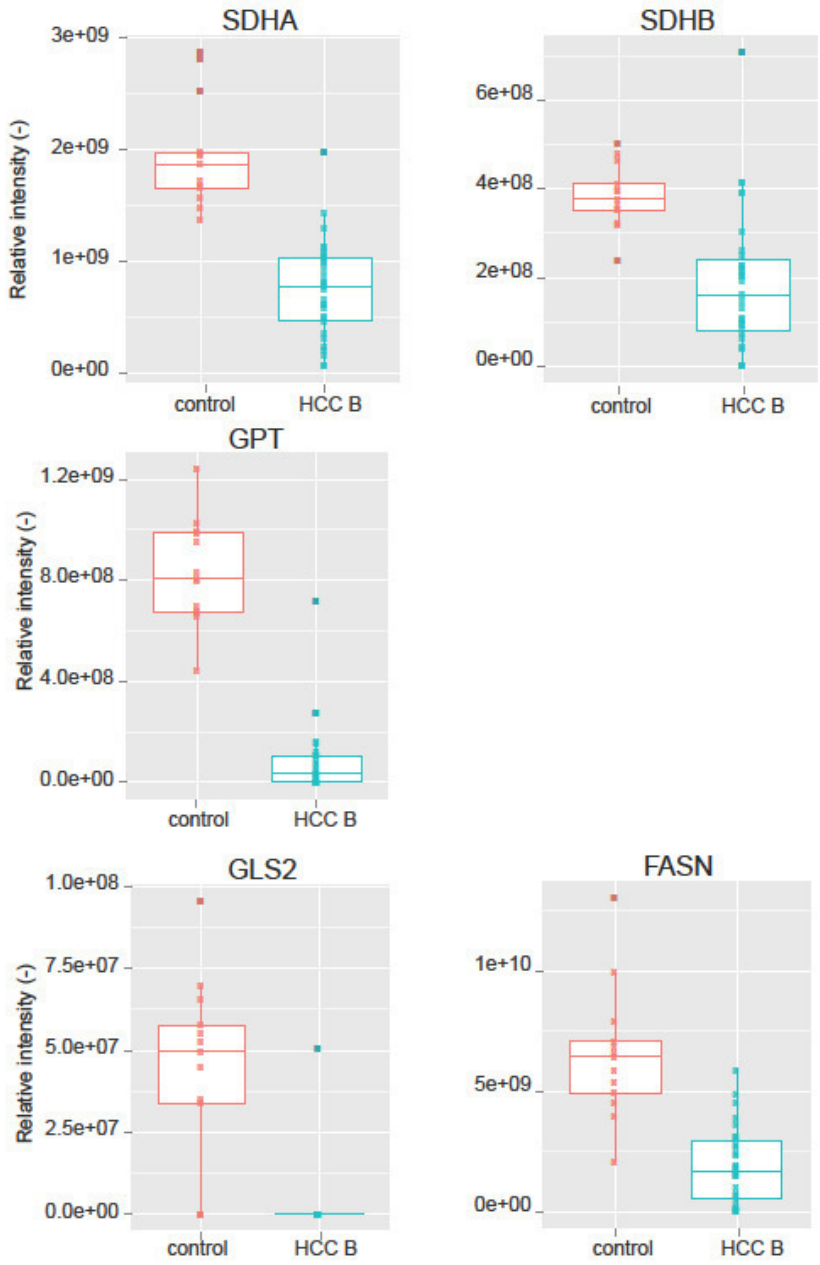


Figure 34. Box plots with the relative intensities of enzymes of TCA cycle and glutaminolysis in human HCC Group B compared to a control group. SHDA was expressed 2.6 times less in HCC (t-

test, $p=1.85*10^{-10}$), and SDHB was 2.1 times less in HCC (t-test, $p=5.06*10^{-5}$). GPT was expressed 10.5 times less in HCC (t-test, $p=1.11*10^{-17}$), and GLS2 was not detected in HCC samples (with the exception of 1 sample). FASN was expressed 3.4 fold more in control samples (t-test, $p=4.46*10^{-19}$).

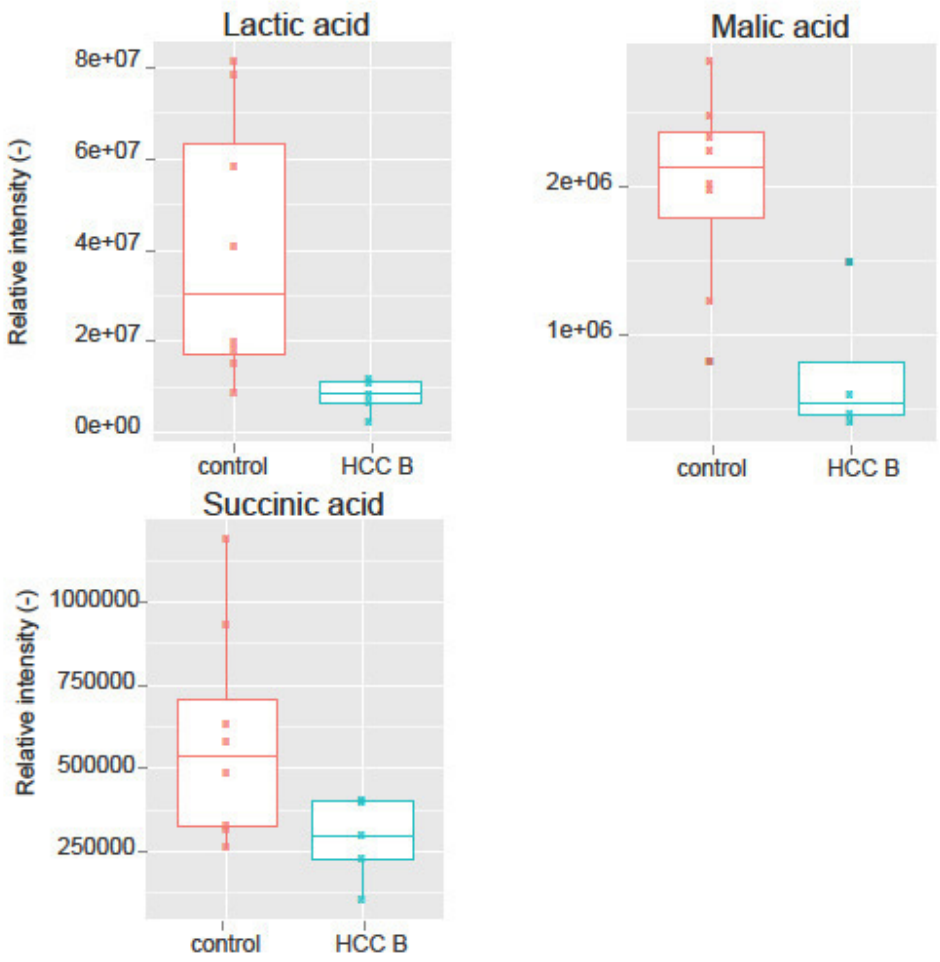


Figure 35. Box plots with the relative intensities of CCM metabolites in human HCC Group B compared to a control group. Lactic acid level was 4.9 lower in HCC (t-test, p=0.018), malic acid level was 2.7 times lower in HCC (t-test, p=0.004), and succinic acid level was 2.1 times lower in HCC (t-test, p=0.04).

Thus, the separation of HCC in group A and B based on hierarchical clustering evaluated with proteomic and metabolomic data revealed different transformations of CMM in Group A and B.

Comparison of CCM changes in human HCC Group A and B

In this section I compare the changes at the proteomic level between Group A and Group B.

An isoform switch was more pronounced in HCC Group B. The average intensity of ALDOA was 3.7 more in Group A and 9.2 times more in Group B compared to control, and ALDOB was 1.5 times less in Group A and 6.3 times less in Group B compared to control. ENO1 was not changed in Group A and it was 1.3 times more in Group B compared to control, ENO3 was 2.3 times less in Group A and was 122 times less in Group B compared to control. PKM was 3.8 higher in Group A, while PKLR was not changed. Group B was characterized by 3 times lower expression of PKLR and 13 times higher expression of PKM compared to control. LDHB was 2.5 times higher in Group A and 6.6 times higher in Group B, LDHD was 1.5 times lower in Group A and 4.7 times lower in Group B compared to control. Pentose phosphate pathway-involved enzyme TKT was 2.9 fold higher in Group A and 5.2 times higher in Group B compared to control. FBP1, which participates in glucose *de novo* synthesis, was 1.4 times lower in Group A and 6.9 times lower in Group B compared to control. Glycogen metabolism: GYS2 was not changed in Group A, and 8.7 times lower in Group B compared to control, PYGL was 1.4 times lower in Group A and 5.4 times lower in Group B. TCA cycle analysis revealed more drastic changes in Group B compared to Group A. ACO1 level was not changed in Group A, and it was 2.7 times lower in Group B compared to control. ACO2 was 1.3 times higher in Group A and 1.3 times lower in Group B. IDH subunits expressed the switch: IDH1 and IDH2 were not changed in Group A, and they were 1.7 and 1.4 times lower in Group B, respectively. IDH3A was 3.5 times higher in Group A and 5.6 times higher in Group B compared to control. GPT, which participates in glutaminolysis along with GLS2, was 1.8 times lower in Group A and 10.5 times lower in Group B. Compared to control GLS2 was 2.2 times lower in Group A, and was not expressed in the Groups B samples.

The approach revealed that Group A of HCC samples underwent less severe dysregulation of CCM proteome pool sizes compared to Group B. The samples were related in a way that reflected the functional changes of liver CCM. Thus, more pronounced and severe changes were typical for HCC cases, which were classified in Group B. Less pronounced functional transitions were found in both, HCC and non-HCC samples. Group A, with moderate changes of CCM, was not fully separated from non-cancerous conditions. Group B contained samples with more drastic, homogeneous and pronounced dysregulation of CCM, which was very consistent among the patients of this group, and fully separated from non-cancerous condition. 66% of HCC samples

were in Group B, which means that they could be detected by focused proteomic analysis. It is important to note, that only Group B contains CTNNB1 mutations. 23% of all samples from Group B contained CTNNB1 mutation.

The metabolomic data obtained from this study showed that overall human metabolome profile remained elusive. Huge diversity within the patients could be explained by the influence of a lot of factors, which the researcher has no influence on: diet, supportive medication, stress level, pre-surgery medication, and others. The optimization options have to be evaluated in further experimental and statistical work.

3. Comparison of mouse and human data

The analysis of changes in CMM in both mouse model in human samples allowed the evaluation of the signature features of HCC in both species. Figure 36 represents the common features changed in both mouse and human HCC. The non-common features are presented in grey for the purposes of simplification.

Proteomic data from the mouse model uncovered a very specific profile of metabolic reprogramming of the liver. In cancer samples glycogen metabolism was impaired via the inhibition of Gys2. Several enzymes of glycolysis showed isoform switches (PGMs, PKs). Gluconeogenesis was downregulated at the level of Pck1 and G6pc. Transport of pyruvate into mitochondria was inhibited at the levels of Mpc1 and Mpc2. 60% of TCA cycle enzymes were downregulated, for example Cs, Sdha, Sdhb and Sdhc. Enzymes responsible for the glutamine uptake to the TCA cycle had lower expression in the HCC samples: Gpt, Gpt2, Gls2 were downregulated at the proteomic level.

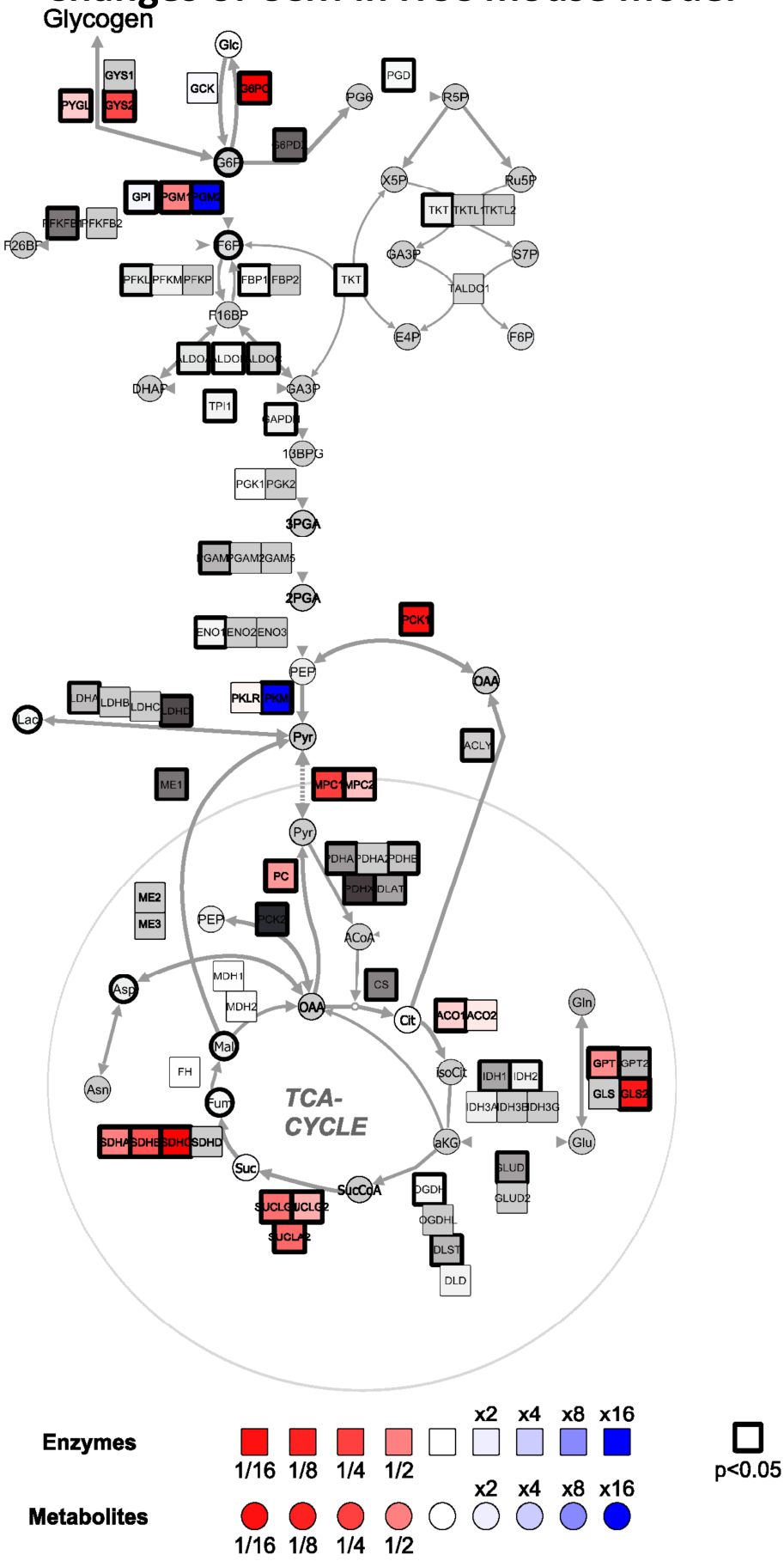
It is important to underline the overall consistency of mice proteomic data. Out of 67 protein used in the CCM analysis, mouse model data showed 52 significantly changed proteins (t-test, 50 with $p<0.005$, and 2 with $p<0.05$). Human samples showed less consistent data. The whole group of human HCC samples included 37 significantly changed proteins compared to the control group (t-test, 19 with $p<0.005$, and 18 with $p<0.05$). Group A HCC samples included 32 significantly changed proteins compared to control (t-test, 10 with $p<0.005$, and 22 with $p<0.05$). Group B HCC samples included 44 significantly changed proteins compared to the control group (t-test, 27 with $p<0.005$, and 17 with $p<0.05$). The higher consistency of the mouse model data could be explained by the same genetic background of the mice, same environmental conditions, and a strictly controlled experimental setup.

Human proteomic, and especially metabolomic, profiles showed drastically lower consistency compared to mouse data. Nevertheless, analysis of the HCC group compared to the control group revealed a signature of tumor-specific metabolic reprogramming (especially in Group B). At the proteomic level glycogen metabolism was downregulated via dysregulation of PYGL and GYS2. Glucose *de novo* synthesis was downregulated at the level of FBP1. TKT and G6PD, which are involved in pentose phosphate pathway, were upregulated in HCC samples (which if different from the mouse model data). Glycolysis was upregulated at the level of hexokinases.

Glycolysis was also characterized by isoform switches at the levels of phosphoglycerate mutases, aldolases, enolases, pyruvate kinases, and lactate dehydrogenases. 65% of TCA cycle enzymes were downregulated, and isocitrate dehydrogenase showed a change of subunit composition. Glutaminolysis was downregulated at the level of GPT and GLS2.

Overall the proteomic data obtained from both mice and humans showed a clear dysregulation of the main energy-producing pathways in HCC, changes in CCM and in the main liver functions. Glycogen metabolism (via PYGL and GYS2), *de novo* glucose synthesis (on various levels), glutamine uptake to the TCA cycle (via GPT, GLS2), and approximately 60% of enzymes of the TCA cycle (SDHs, SUCLs, ACO etc.), and transport of pyruvate to mitochondria (MPCs) were downregulated in HCC. Additionally the proteome data display one of the most common features of the cancer cells: metabolic reprogramming. An isoform switch, which is one of the fundamental mechanisms in reprogramming, was demonstrated in this work. Both mice and humans with HCC revealed isoform switches at the level of phosphoglycerate mutases and pyruvate kinases. Despite the diversity of genetic background of human samples the proteome described HCC signature was very robust in both bioptate and surgery material. The key features were found in both mouse and human, showing a universal metabolic HCC fingerprint. This allowed me to confidently conclude, that the ASV-B mouse model appeared to be sufficiently similar to the human clinical presentation that it is useful for further study of the disease.

Changes of CCM in HCC mouse model



Changes of CCM in HCC human samples (Group B)

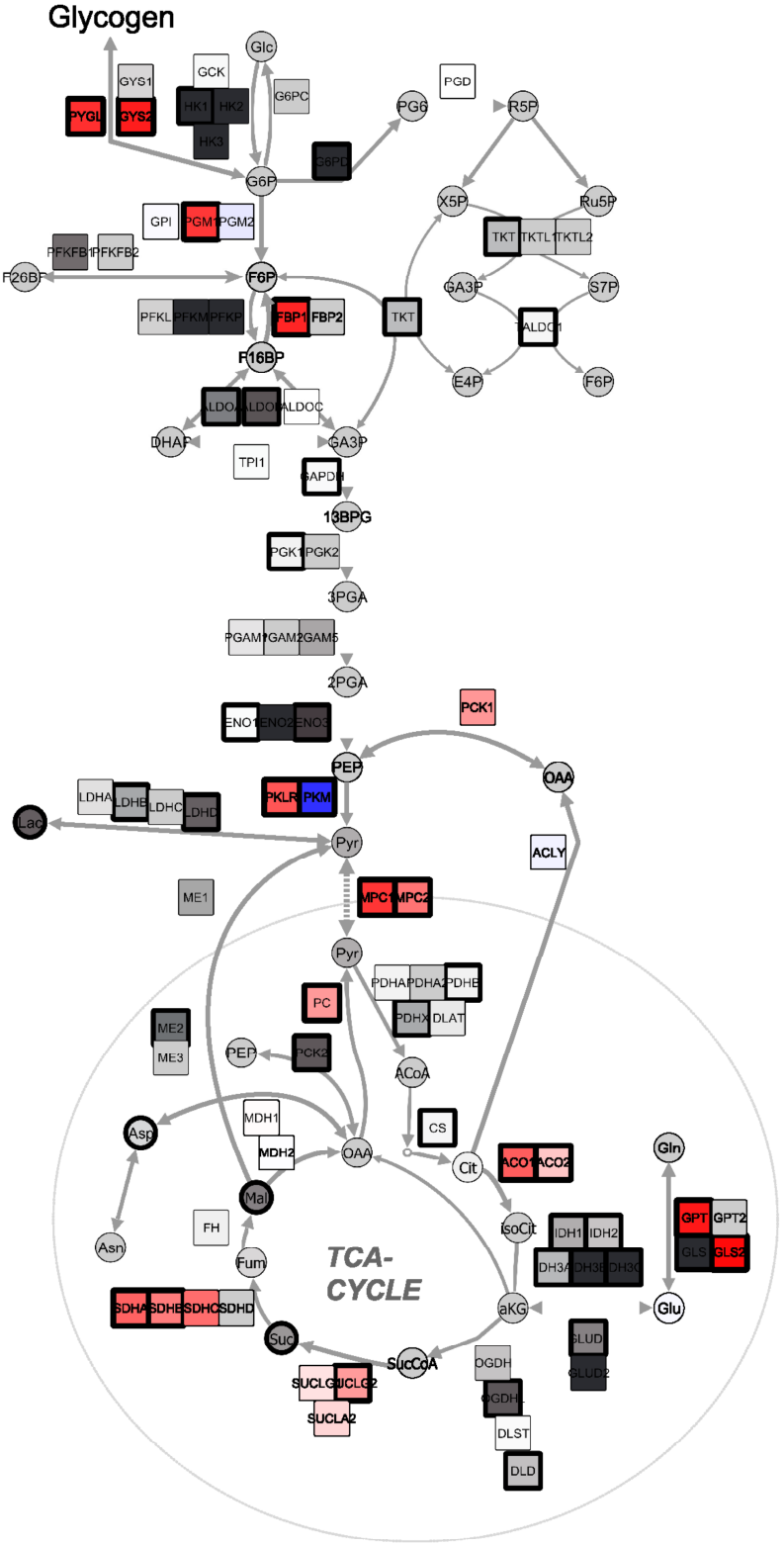


Figure 36. The common features changed in both mouse and human HCC compared to a respective control group. The non-common features are presented in grey for the purposes of simplification. The metabolites are visualized with circles, the proteins are visualized with squares, bold frame indicates the changes with $p<0.05$ (t-test), grey colour of the metabolite (circle) or protein (square) at the scheme indicates that data regarding certain enzyme or metabolite is not available.

4. Serum data

I obtained the proteomic data to evaluate changes in human serum samples. The cohort included 5 HCC patients, from whom the serum was taken prior to and from 5 to 7 days after liver resection surgery. 5 patients with either hemangioma[1] or hepatic cyst[2] formed the control group.

The proteins from human serum samples were extracted and analyzed on a Q Exactive HF. 460 proteins were identified. An abundance index was calculated for each protein: the intensity of a protein was divided by molecular mass of the protein and further each value in the column was divided by the mean intensity value of the column. This normalization performed on the data set is described in the paper of Mastrobuoni et al. 2012 (Mastrobuoni, et al., 2012). Imputation of missing values for both matrices (serum and bioptate material) was performed by replacing the missing values with a constant value (50% of the minimum non-0 value in the matrix). Here I applied the imputation, because for clinical purposes it is better to have imputed values of fold changes. The implication of data imputation revealed 15 proteins (Table 4) significantly different in serum from control samples compared to HCC.

[1] The hepatic hemangioma (haemangioma) is a benign vascular lesion. Histologically, it is a mesenchymal lesion consisting of blood-filled of various sizes vascular cavities, surrounded by a layer of endothelial cells, supported by a fibrous connective tissue. (Klotz, et al., 2013)

[2] A cyst is a closed cavity with a distinct membrane and division compared to the nearby tissue. Hepatic cysts are classically divided into parasitic and non-parasitic types, with the latter being the most prevalent worldwide. (Maruyama, et al., 2013)

Table 4. Significantly changed proteins in serum samples: HCC compared to control group.

Protein IDs	Protein names	tumor/ctrl	t-test	Gene names
P08697	Alpha-2-antiplasmin	0.572	0.005	SERPINF2
P01613	Ig kappa chain V-I region Ni	2970.849	0.005	
P35858	Insulin-like growth factor-binding protein complex acid labile subunit	0.348	0.005	IGFALS
P36980	Complement factor H-related protein 2	692.345	0.010	CFHR2
P01857	Ig gamma-1 chain C region	1.390	0.015	IGHG1
P26927;Q2TV78	Hepatocyte growth factor-like protein; Putative macrophage stimulating 1-like protein	81.170	0.016	MST1;MST1L
P06681	Complement C2	1.375	0.022	C2
Q15848	Adiponectin	547.195	0.026	ADIPOQ
Q92820	Gamma-glutamyl hydrolase	5.406	0.033	GGH
P61769	Beta-2-microglobulin	2551.626	0.035	B2M
P27169	Serum paraoxonase/arylesterase 1	0.497	0.040	PON1
P61626	Lysozyme C	193.375	0.041	LYZ
P19652	Alpha-1-acid glycoprotein 2	1.706	0.041	ORM2
P48740	Mannan-binding lectin serine protease 1	47.672	0.041	MASP1
P00738	Haptoglobin	3.627	0.045	HP

806 proteins were found to be significantly changed in HCC bioptate material compared to control. Venn diagram (Figure 37) shows the amount of common proteins among significantly different proteins in bioptate and serum group: B2M, PON1, HP (Table 5) (VIB/UGent, 2017).

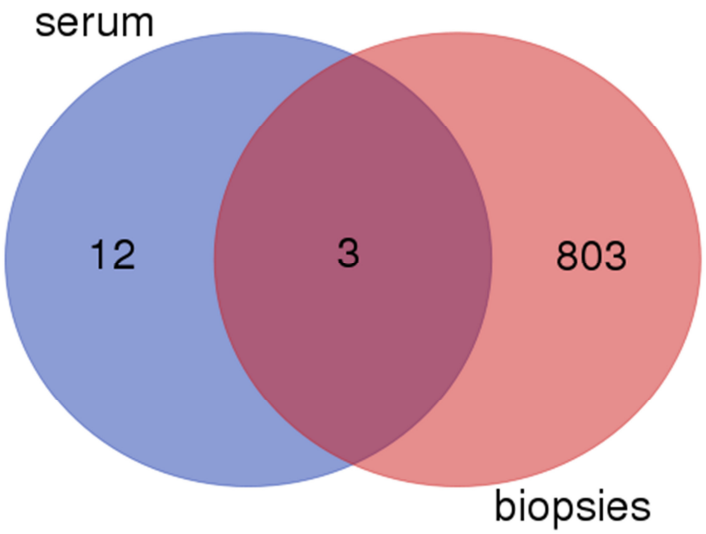


Figure 37. The amount of common proteins among significantly different proteins in bioptate and serum. These proteins are B2M, PON1, HP.

Table 5. Proteins, which are significantly differentially expressed in human serum and biopsy HCC material.

Gene names	Protein IDs	Protein names	Serum		Biopsies	
			tumor/ctrl	t-test	tumor/ctrl	t-test
B2M	P61769	Beta-2-microglobulin	2551.626	0.035	2.654	0.008
PON1	P27169	Serum paraoxonase/arylesterase 1	0.497	0.040	0.560	0.0001
HP	P00738	Haptoglobin	3.627	0.045	5.955	0.001

Significantly different B2M, PON1, and HP changes in HCC were reflected in the serum, which revealed new opportunities for specific, sensitive, and non-invasive diagnostic tool of HCC. While three of the suggested diagnostic markers need to be validated further on larger cohorts as a potential diagnostic tool, all of the 15 proteins mentioned above are to be investigated for the prognostic marker efficacy.

Discussion

The following section focuses on various sides of the biological relevance and medical implications of my results. I also approach the challenges of present work. The cohort formation was one of the main issues I dealt with from the beginning of the project, as well as tissue acquisition and any possible bias it may cause. Further I discuss the interpretation of multiomics data and its application to CCM research. The last part of Discussion is devoted to the analysis of serum from HCC patients.

Cohort and material

A significant part of this work was devoted to the cohort development. Several issues had to be resolved during the collection period in order to account for the heterogeneity among patients. Since liver is one of main metabolic organs, it was essential to try to avoid patents with metabolic disorders, such as diabetes and obesity. The pathological mechanism of NAFLD and fibrosis or cirrhosis development with an obesity[1] as a background condition is well-known. However, the further mechanism of progression to HCC is not studied properly for both groups of patients: diabetic and obese and patients without these conditions. Only 7 patients from our cohort had obesity, 5 patients had diabetes.

Most of researchers have a tendency to combine etiologies in one study, without separating various causes and risk factors. This could lead to the certain limitations in the discovery of specific markers and on cause-specific changes in CCM. For examples, Tan et al. (Tan, et al., 2014) published the biomarker panel (which included HPS70, TKT and other proteins) for metastatic HCC, combining all associated diseases in one group of HCC samples. The cohort includes HBV, HCV, and virus-free tissues. This approach tends to be cohort-biased and non-specific. Similar problems were faced by Orimo et al. (Orimo, et al., 2008) in their proteomic study of HCC, looking for a prognostic marker in HCC. The cohort the authors worked with included HBV, HCV, HBV and HCV, and virus-free patients. While the suggested marker EB1 (APC-binding protein EB1) showed the significant results, the fundamental mechanisms of the described changes remain elusive.

Another problem researchers deal with is the liver tissue material source. In vast majority of cases surgery material is used (Beyoglu, et al., 2013; Orimo, et al., 2008; Tan, et al., 2014), which is valid for proteomic analysis, but, as I showed in my work (Figure 13), has certain limitations in revealing the *in situ* metabolic picture. For example, Beyoğlu et al. state, that no differences in

[1] '2 hit hypothesis' by (Dowman, Tomlinson, & Newsome, 2010)

glucose, lactate, malate, concentrations were found, suggesting that the CTNNB1 activates Wnt/β-catenin pathway by mutation in groups G5 and G6 shows no specific remodeling of metabolism (Beyoglu, et al., 2013). This observation could possibly be bound to the usage of exclusively surgically-delivered liver tissue for metabolomic analysis. The opportunity to involve biopsy material in the further metabolic research would eliminate the obstacles described above.

The cohort from the present work contained both surgery and biopsy tissue. Control samples were acquired exclusively by biopsy, while nearly 75% of HCC samples come from surgery material. The difference of the source of material could explain the high number of significantly changed proteins in HCC group of samples compared to non-HCC (Results, Human cohort, Proteomics, Proteome overview, Figure 11). When all the non-HCC samples come from biopsies, and most of the HCC samples from surgery, the result could be acquisition method-bias. However, further proteome-based analysis (Results, Human cohort, Multiomics picture, Classification, Figure 15) showed that surgery and biopsy samples did not cluster according to the source of tissue. This observation proved that the chosen method of separating HCC on Groups A and B according to the proteomic changes of CCM was not affected by the choice of tissue source. This statement is fair for the proteomic data, and requires further validation for other types of analyses. Nonetheless, in future work I would prefer to work with biopsy tissue, because it more clearly represents the *in situ* picture of CCM changes.

CCM changes

Mouse model overview

Proteomic data from the ASV-B mouse model showed a specific metabolic reprogramming of damaged liver (Figure 36). In cancer liver tissue glycogen metabolism was impaired, enzymes of glycolysis revealed isoform switches at the level of phosphoglycerate mutases and pyruvate kinases, gluconeogenesis was downregulated, as well as the pentose phosphate pathway. Transport of pyruvate to mitochondria was inhibited, 60% of TCA cycle enzymes were downregulated, and glutaminolysis was downregulated at the proteomic level. At the metabolic level the upregulation of glycolysis was detected.

Human cohort overview

The signature of tumor-specific metabolic reprogramming also was found in the proteomic profile of human samples (Figure 36). At the proteomic level glycogen metabolism was downregulated, glucose *de novo* synthesis was downregulated, and pentose phosphate pathway was upregulated in HCC (which is different from the mouse model data). Glycolysis was upregulated at the level of hexokinases. Glycolysis was also characterized by an isoform switch at

the levels of phosphoglycerate mutases, aldolases, enolases, pyruvate kinases, and lactate dehydrogenases. The inhibition of pyruvate transport to mitochondria was observed at the proteomic level, to be more specific mitochondrial pyruvate carriers 1 and 2 both were expressed less in HCC compared to control. 65% of the enzymes of the TCA cycle were downregulated, and isocitrate dehydrogenase showed a switch between subunits. Glutaminolysis in HCC sampleswas downregulated. Due to the tremendous heterogeneity of the phenotypes of human patients, metabolic data was not consistent over the entire cohort. However, a tendency of the metabolites of lower glycolysis (pyruvate, lactate), and metabolites of the TCA cycle to decrease in HCC, could be observed.

The approach revealed that Group A of HCC samples underwent less severe dysregulation of CCM proteome pool sizes compared to Group B. The samples were related in a way that reflected the functional changes of liver CCM. Thus, more pronounced and severe changes were typical for HCC cases, which were clustered in Group B. Less pronounced functional transitions were found in both HCC and non-HCC samples. Group A, with moderate changes of CCM, was not fully separated from non-cancerous conditions. Group B contained samples with more drastic, homogeneous and pronounced dysregulation of CCM. Described changes were very consistent among the patients of this group and fully separated from non-cancerous conditions. Interestingly, the human cohort included 2 patients with recurrent HCC (74, 84), and both of them fit in Group B. This may indicate that recurrent HCC tumors reveal more drastic changes of CCM, or that tumors with more drastic changes of CCM are prone to recidive. However, this statement needs to be validated on larger cohorts of patients. Additional information about patients, such as AFP level, AST and ALT levels, and, most importantly, at least a 5-year follow up of outcome, would allow the evaluation the prognostic possibilities of this classification. Testing for alcohol consumption blood biomarkers[1] would complement to the cause evaluation and outcome of HCC, due to patients' resistance to admit alcohol addiction (less than 10% of patients claimed alcohol abuse).

The phenomenon of lower level of lactic acid was detected in human HCC samples (Figure 17, 23, 29). This observation does not fit to the Warburg effect. Under normal nutritional conditions I would expect to detect the Warburg effect at the metabolic level. However, fasting

[1] However, the specificity of most currently suggested blood biomarkers of alcohol consumption is questionable. (Allen & Wilson, 2003; Peterson, 2004)

caused the overall decrease in nearly all of detected metabolites (Figure 29), due to the high dependency of liver tumor on glucose. Thus, supposedly, by the time of material acquisition glucose was used, and levels of pyruvic acid and lactic acid dropped.

Comparison of mouse model and human cohort

Obtained proteome and metabolome data from both mice and humans showed the clear dysregulation of the main energy-producing pathways in HCC, as shown at the scheme on Figure 37. Glycogen metabolism, *de novo* glucose synthesis, glutamine uptake to the Krebs cycle, and other processes were less active in HCC. The proteome and metabolome data taken together displayed one of the most common features of the cancer cells: the metabolic reprogramming. Isoform switch is one of the fundamental mechanisms in reprogramming, was demonstrated in this work. These observations revealed the universal, for both mouse and human species, mechanism of CCM changes in HCC. Due to the similarities of basic processes in HCC development, this study opened the possibilities for anti-HCC drug testing on animal model. ASV-B mouse model allows evaluating the exact CCM changes and medication influence on them.

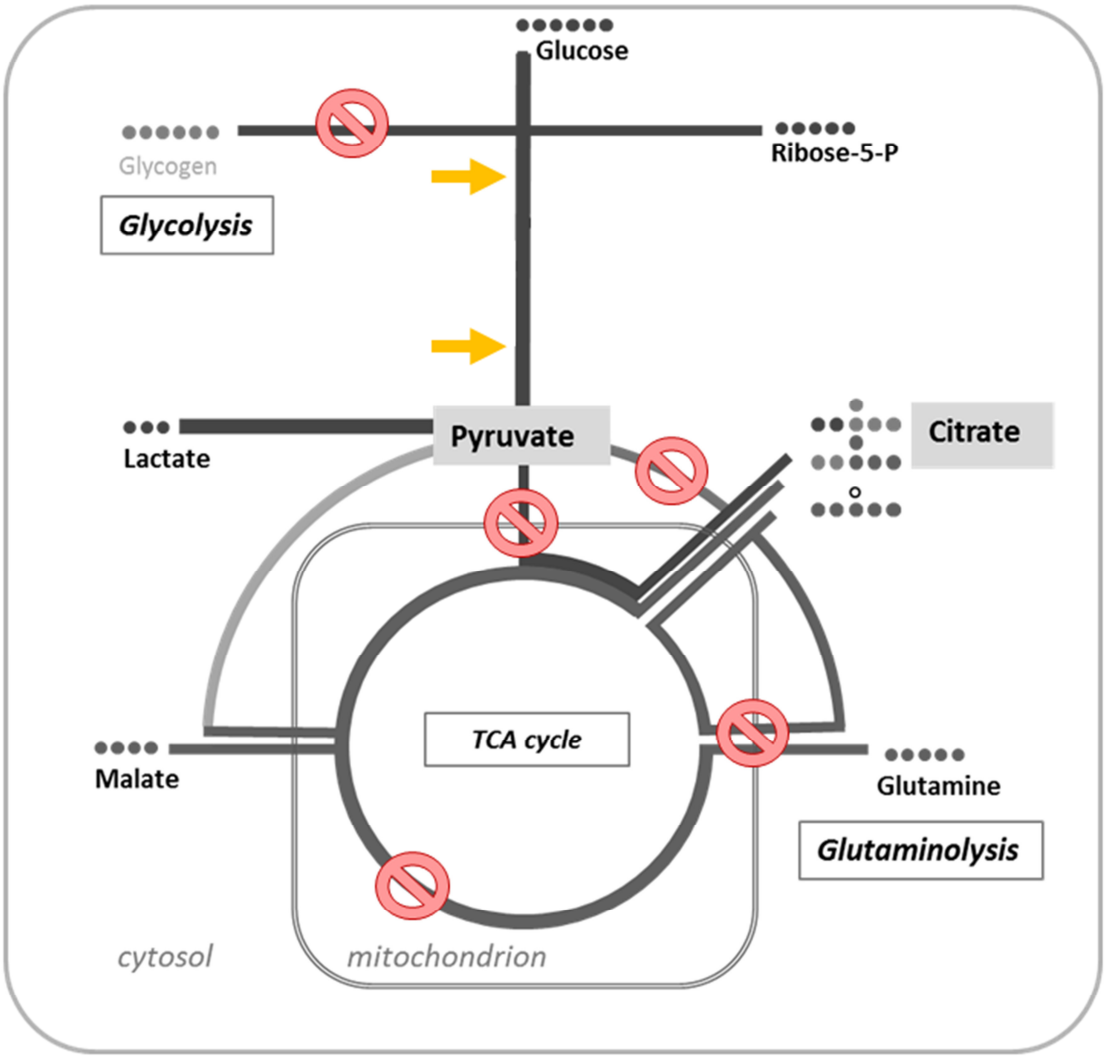


Figure 37. The scheme visualizing the similarities of CCM changes in mouse HCC model and human HCC samples compared to control. The scheme represents common features in HCC metabolism for both mice and human. Cancer cells lost the ability to break down and synthesize glycogen, to perform gluconeogenesis, to transport pyruvate to mitochondria, to take amino acids to the TCA cycle. Approximately 60% of TCA cycle enzymes were downregulated. An isoform switch was observed at the level of phosphoglycerate mutases and pyruvate kinases. Red circles indicate the downregulated areas, yellow arrows point out areas with an isoform switches.

This study demonstrates that the proteome analysis of the bioptate material is a strong and sufficient molecular diagnostic tool for research in cancer: the proteomic analysis of liver material allows a clear distinction of tumor samples from non-tumor samples, and tracks the level of the

disease progression. However the described molecular tools are applicable to material obtained by invasive techniques, such as biopsies and surgeries. Therefore we applied the developed proteomics technology and supervised statistical methods on serum taken from human patients, in order to evaluate the non-invasive diagnostic tools.

Mutations

Targeted genome sequencing was used in order to detect cancer-related somatic mutations in human samples. 1000 Genomes Project (The Genomes Project, 2015) database was used to evaluate the SNPs and non-harmful mutations. Due to the fact that we only worked with pathological tissue, there was no possibility to eliminate person-specific SNPs from analysis, thus we had to rely on 1000 Genome Project. The final data set from Project contains data for 2,504 presumably healthy individuals from 26 populations, which allows to estimate the level of harmfulness of one mutation or another. However, the presence of certain mutation in presumably healthy individual does not necessarily mean that this mutation is not harmful, and that the individual will not develop a disease later. A disease could be on early developing stage, thus not affecting the picture of blood tests of subjective feeling of well-being. Another issue with presumably harmless SNPs that the Project does not consider the combination of those SNPs. Possibly the combination of several 'harmless' SNPs can cause the disease. Another limitation of this work was the application of panel sequencing. The panel sequencing is restricted to the known set of mutations, unlike the whole genome sequencing. In our cohort were patients with no mutations detected, which does not mean that individuals had no mutations. Nonetheless, various mutations were detected, identified in confirmed in the cohort.

CTNNB1

Several CTNNB1 mutations are specific for tumor samples and were detected in 15% of HCC. 23% of Group B samples contained CTNNB1 mutations: c.A121G, p.T41A; c.T133C, p.S45P; c.C134T, p.S45F; c.C110T, p.S37F; c.C98T, p.S33F, all of which were located in exon 3 (Gamallo, et al., 1999).

CTNNB1 (c.C134T, p.S45F) mutation has been shown in patients with desmoid tumors (this type of tumors originates from fibroblasts) (Colombo, et al., 2013; Mullen, et al., 2013; Nishida, et al., 2016). Nishida et al. and Mullen at al. also found CTNNB1 (c.A121G, p.T41A) mutation in desmoid tumors. Doyen et al. show the CTNNB1 (c.A121G, p.T41A) and CTNNB1 (c.T133C, p.S45P) mutations and suggest that a certain CTNNB1 mutations are necessary to provide the tumor with

the growth advantage (Doyen, et al., 2016). Kolble et al. describes the CTNNB1 (c.T133C, p.S45P) mutation being present in colorectal cancer and suggests that this specific mutation if restricted to a hepatic metastasis (Kolble, et al., 2000). CTNNB1 (c.C98T, p.S33F) mutation can be involved in the development of colorectal cancer (Alomar, et al., 2016).

Studies of Gamallo et al., Kobayashi et al. and Moreno-Bueno et al. report the CTNNB1 (c.A121G, p.T41A) and CTNNB1 (c.C110T, p.S37F) mutations existence in ovarian and endometrial carcinomas (Gamallo, et al., 1999; Kobayashi, Sagae, Nishioka, Tokino, & Kudo, 1999; Moreno-Bueno, et al., 2001). Kobayashi et al. suggest, that each mutation of CTNNB1 with the serine/threonine substitution results in the downregulation of beta-catenin through phosphorylation by GSK-3 beta kinase. The work of Palacios and Gamallo also evaluated the role of mutation of CTNNB1 on exon 3, confirming the (c.A121G, p.T41A) and (c.C110T, p.S37F) mutations, implicating CTNNB1 mutations in ovarian malignant transformation (Palacios & Gamallo, 1998).

None of the above mentioned mutations were previously described in HCC, however the suggested by several authors mechanism of malignisation of ovarian and endometric tissues could be applicable to the liver cancer development. Alternatively, considering the desmoid tumors publications (Colombo, et al., 2013; Mullen, et al., 2013; Nishida, et al., 2016), the origin of the tumors, diagnosed as hepatocellular, can be questionable. Here we excluded the ovarian cells origin because 5 out of 6 patients with CTNNB1 mutations are males (information about 1 patient is not available).

However, several CTNNB1 mutations are specific for tumor samples and were detected in 15% of HCC. This observation can indicate that CTNNB1 mutation is not a necessary but sufficient factor for the CCM changes, and, when present, affects metabolism in a drastic way, which has an impact on major liver functions. Thus, the above mentioned CTNNB1 mutations should further be validated on larger cohorts, for potential use as a prognostic biomarker of survival or metastatic status.

Overall, CTNNB1 mutations present in current cohort correlated with more drastic changes in CCM, which were be detected by proteome analysis. CTNNB1 mutations should further be validated on larger cohorts for potential use as a prognostic biomarker of survival or metastatic status. Additionally, the mechanistic studies of the influence of CTNNB1 mutations on the changes of malignancy potential and invasiveness can be performed on liver and HCC cell lines.

TP53

The TP53 (c.216delC, p.V73fs) mutation was frequently met in the full cohort. This specific mutation is described by Mouradov et al. in the KM12 colon cancer cell line (Mouradov, et al., 2014). This paper also present the TP53 (c.902delC, p.P301fs), which was met in the cohort twice, in another colon cancer cell line – HCA-7. In addition, TP53 (c902delC, P301fs) was confirmed to be present in a colon cancer patient (Giannakis, et al., 2014).

PIK3CA

Janku et al. reported PIK3CA (c.A3140G, p.H1047R) mutation to be found in colorectal cancer patients (Janku, et al., 2012). Reiner et al. studied the cohort of 50 patient with HCC and found PIK3CA (c.A3140G, p.H1047R) mutation if 2% of them. Their results suggest that somatic PIK3CA mutations play role in the frequent activation of the PI3K/AKT pathway in carcinomas of the biliary tract and liver (Riener, Bawohl, Clavien, & Jochum, 2008).

TP53, CTNNB1, and PIK3CA mutations are known in the literature to be associated with HCC in HBV and HCV contaminated patients (Tornesello, et al., 2013). However, in our cohort these mutations are present not exclusively in patients with HBV or HCV.

The other mutations represented in Table 2 show no solid evidence in the literature to be involved in HCC development or functioning processes. The mutations were detected in the samples with fatty liver, hepatitis B and C, fibrotic liver, cirrhosis, and HCC. However, no specific mutations, apart from CTNNB1, were confirmed to be present only in HCC. No clear distinction between cancer and precancerous conditions could be achieved exclusively by the mutation analysis.

Overall my analysis unravels that the mutations in HCC are more wide spread than expected.

Serum analysis

The developed shotgun proteomics technology was applied to serum samples taken from human patients, in order to evaluate the non-invasive diagnostic tools. Serum analysis revealed 460 proteins (mostly components of complement, immunoglobulins and coagulation factors), 15 of which were differentially expressed in tumor group compared to the group of samples without tumor. 806 proteins were found to be significantly changed in HCC bioptate material compared to control. The overlap of both serum and bioptate cohort uncovered 3 proteins: B2M, PON1, HP.

B2M - Beta-2-microglobulin – component of the class I major histocompatibility complex (MHC). B2M is involved in the presentation of peptide antigens to the immune system

(UniProtconsortium, 2017). Ouda et al. detect approximately 5-fold elevation of B2M in plasma samples from patients with HCV-infected HCC compared to control (Ouda, Khairy, Sorour, & Mikhail, 2015).

PON1 - Serum paraoxonase/arylesterase 1 – hydrolyzes the toxic metabolites, is capable of hydrolyzing a broad spectrum of organophosphate substrates and lactones. Mediates an enzymatic protection of low density lipoproteins against oxidative modification and atheroma formation (UniProtconsortium, 2017). Elevation of PON1 in plasma was detected in HCV-infected patients with alcohol abuse (Ferrin, et al., 2015).

HP - Haptoglobin – captures free plasma hemoglobin to allow hepatic recycling of heme iron and to prevent kidney damage. Haptoglobin also acts as an antioxidant, has antibacterial activity and affects many aspects of the acute phase response. Hemoglobin/haptoglobin complexes are rapidly cleared by the macrophage scavenger receptor expressed on the surface of liver Kuepfer cells, using a pathway of an endocytic lysosomal degradation (UniProtconsortium, 2017). Altered levels of fucosylated haptoglobin in serum is used in diagnostics of HCC (Mondal, Saroha, Bose, & Chatterjee, 2016; S. Shang, Qin, Li, Zhang, & Liu, 2016; Tawara, et al., 2016).

B2M, PON1, HP may have a diagnostic significance. Serum proteome profile reflected proteome changes in damaged liver samples obtained from selected patients. These markers potentially could be detected in serum, avoiding invasive and potentially harmful biopsy obtaining procedure. However additional studies on large cohorts have to be performed in order to evaluate diagnostics and/or prognostic perspectives.

Conclusions and outlook

In the present work I have shown on mouse model and human samples that the metabolism of hepatocellular carcinoma is reprogrammed.

ASV-B HCC mouse model showed a distinct reprogramming of central carbon metabolism at the proteomic level. Liver damaged with cancer revealed the impaired glycogen metabolism at the level of glycogen synthase 2 (Gys2). The gluconeogenesis was impaired at the level of phosphoenolpyruvate carboxykinase 1 (Pck1) and glucose-6-phosphatase catalytic-subunit (G6pc). Pyruvate transport to mitochondria was downregulated via the inhibition of pyruvate carriers 1 and 2 (Mpc 1, 2). Most of the TCA cycle proteins were downregulated, as well as the glutamine uptake to TCA cycle at the level of glutamate pyruvate transaminase (Gpt), glutamate pyruvate transaminase 2 (Gpt2), and glutaminase 2 (Gls2). The isoform switch was shown at the level of phosphoglycerate mutases (Pgm) and pyruvate kinases (Pk)). At the metabolic level the higher amounts of glucose-6-phosphate, fructose-6-phosphate and lactic acid were detected. The metabolites of TCA cycle were characterized by higher levels in HCC. Fumaric and malic acid levels were higher in cancer compared to control.

Human samples revealed a similar signature of tumor-specific reprogramming of CCM. At the proteomic level, glycogen metabolism was inhibited at the level of glycogen phosphorylase L (PYGL) and glycogen synthase 2 (GYS2). Gluconeogenesis was dysregulated at the level of fructose-1,6-bisphosphatase 1 (FBP1). The upregulation of hexokinases (HK) characterizes an increased glycolysis, as well as an isoform switch at the levels of phosphoglycerate mutases (PGM), aldolases (ALDO), enolases (ENO), pyruvate kinases (PK), and lactate dehydrogenases (LDH). 65% of the proteins of the TCA cycle were downregulated, and isocitrate dehydrogenase (IDH) showed a switch between subunits. Glutamine uptake to TCA cycle was inhibited at the level of glutamate pyruvate transaminase (GPT) and glutaminase 2 (GLS2). Metabolomic picture was mostly inconclusive, possibly due to the incredible diversity of genotypes and phenotypes in human cohort. This diversity may come from different environmental conditions, diets, and comorbidities of patients in cohort.

In sum, proteomic data from both mice and humans showed a clear downregulation of the main energy-producing pathways in HCC. Glycogen metabolism (via PYGL and GYS2), gluconeogenesis (on various levels), glutaminolysis (via GPT, GLS2), and 60-65% of enzymes of TCA

cycle (SDHs, SUCLs, ACO etc.), and transport of pyruvate to mitochondria (MPCs) were downregulated in HCC. The proteomic data obtained in this work demonstrated one of the most common features of cancer cells: metabolic reprogramming. An isoform switch - one of the fundamental mechanisms in cancer reprogramming - was demonstrated in present study. Isoform switches at the level of phosphoglycerate mutases and pyruvate kinases were revealed in both mouse model and human samples. Even though the genetic backgrounds of human samples were extremely heterogeneous, the proteomic HCC signature was very robust and distinct in both bioptate and surgery material. The key features were found in both mouse and human, showing a universal metabolic HCC fingerprint. This observation confirmed that the chosen ASV-B mouse model was perfectly suitable to study HCC metabolism.

The developed shotgun proteomics technology was used to analyze the serum from human patients, in order to evaluate the options of development the non-invasive diagnostic tool. Out of 460 detected proteins, 15 were differentially expressed in tumor serum samples compared control group. The bioptate material analysis revealed 806 significantly changed proteins in HCC compared to control group. However, only 3 proteins were detected to be significantly changed in both serum and bioptate cohorts: B2M, PON1, HP. These proteins may have a diagnostic significance, however, this statement has to be confirmed on larger cohorts in further studies. Additionally, this finding suggests that serum does in part reflects the proteomic changes in liver, yet this observation has to be validated in larger cohorts as well. Serum proteome profile, allegedly, reflects proteome changes in damaged liver.

This study confirmed that HCC metabolism might serve as therapy target due to its robust picture, which was independent on genetic background and mutations heterogeneity, and environmental conditions. Further molecular analyses of central carbon metabolism changes can improve the non-invasive diagnostics of HCC, and may shed the light on prognostic markers of outcome and possible treatment success.

Bibliography

Abelev, G. I., Perova, S. D., Khramkova, N. I., Postnikova, Z. A., & Irlin, I. S. (1963). Production of embryonal alpha-globulin by transplantable mouse hepatomas. *Transplantation, 1*, 174-180.

Allen, J. P., & Wilson, V. B. (2003). *Assessing alcohol problems. A Guide for Clinicians and Researchers* (2 ed.): National Institute on Alcohol Abuse and Alcoholism, NIH.

Alomar, S. Y., Mansour, L., Abuderman, A., Alkhuriji, A., Arafah, M., Alwasel, S., Harrath, A. H., Almutairi, M., Trayhyrn, P., & Dar, J. A. (2016). beta-Catenin accumulation and S33F mutation of CTNNB1 gene in colorectal cancer in Saudi Arabia. *Pol J Pathol, 67*, 156-162.

Altekruse, S. F., McGlynn, K. A., & Reichman, M. E. (2009). Hepatocellular Carcinoma Incidence, Mortality, and Survival Trends in the United States From 1975 to 2005. *Journal of Clinical Oncology, 27*, 1485-1491.

Andreana, L., Isgro, G., Marelli, L., Davies, N., Yu, D., Navalkissoor, S., & Burroughs, A. K. (2012). Treatment of hepatocellular carcinoma (HCC) by intra-arterial infusion of radio-emitter compounds: trans-arterial radio-embolisation of HCC. *Cancer Treat Rev, 38*, 641-649.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nat Genet, 25*, 25-29.

Bergstrand, C. G., & Czar, B. (1956). Demonstration of a new protein fraction in serum from the human fetus. *Scand J Clin Lab Invest, 8*, 174.

Berth, M., Moser, F. M., Kolbe, M., & Bernhardt, J. (2007). The state of the art in the analysis of two-dimensional gel electrophoresis images. *Appl Microbiol Biotechnol, 76*, 1223-1243.

Bertino, G., Di Carlo, I., Ardiri, A., Calvagno, G. S., Demma, S., Malaguarnera, G., Bertino, N., Malaguarnera, M., Toro, A., & Malaguarnera, M. (2013). Systemic therapies in hepatocellular carcinoma: present and future. *Future Oncol, 9*, 1533-1548.

Beyoglu, D., Imbeaud, S., Maurhofer, O., Bioulac-Sage, P., Zucman-Rossi, J., Dufour, J. F., & Idle, J. R. (2013). Tissue metabolomics of hepatocellular carcinoma: tumor energy metabolism and the role of transcriptomic classification. *Hepatology, 58*, 229-238.

Bielow, C., Mastrobuoni, G., & Kempa, S. (2016). Proteomics Quality Control: Quality Control Software for MaxQuant Results. *J Proteome Res, 15*, 777-787.

Boiteux, A., & Hess, B. (1981). Design of glycolysis. *Philos Trans R Soc Lond B Biol Sci, 293*, 5-22.

Borentain, P., Gerolami, V., Ananian, P., Garcia, S., Noundou, A., Botta-Fridlund, D., Le Treut, Y. P., Berge-Lefranc, J. L., & Gerolami, R. (2007). DNA-repair and carcinogen-metabolising enzymes genetic polymorphisms as an independent risk factor for hepatocellular carcinoma in Caucasian liver-transplanted patients. *Eur J Cancer, 43*, 2479-2486.

Bosman, F. T., Carneiro, F., Hruban, R. H., & Theise, N. D. (2010). *WHO Classification of Tumours of the Digestive System* (4 ed. Vol. 3).

Boyault, S., Rickman, D. S., de Reynies, A., Balabaud, C., Rebouissou, S., Jeannot, E., Herault, A., Saric, J., Belghiti, J., Franco, D., Bioulac-Sage, P., Laurent-Puig, P., & Zucman-Rossi, J. (2007). Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology, 45*, 42-52.

Burnette, W. N. (1981). "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem, 112*, 195-203.

Chang, C. C., Chen, Y. J., Huang, T. H., Chen, C. H., Kuo, F. Y., Eng, H. L., Yong, C. C., Liu, Y. W., Lin, T. L., Li, W. F., Lin, Y. H., Lin, C. C., Wang, C. C., & Chen, C. L. (2017). Living Donor Liver

Transplantation for Combined Hepatocellular Carcinoma and Cholangiocarcinoma: Experience of a Single Center. *Ann Transplant, 22*, 115-120.

Chiesa, R., Donato, F., Tagger, A., Favret, M., Ribero, M. L., Nardi, G., Gelatti, U., Bucella, E., Tomasi, E., Portolani, N., Bonetti, M., Bettini, L., Pelizzari, G., Salmi, A., Savio, A., Garatti, M., & Callea, F. (2000). Etiology of hepatocellular carcinoma in Italian patients with and without cirrhosis. *Cancer Epidemiol Biomarkers Prev, 9*, 213-216.

Clasen, S., Rempp, H., Hoffmann, R., Graf, H., Pereira, P. L., & Claussen, C. D. (2014). Image-guided radiofrequency ablation of hepatocellular carcinoma (HCC): is MR guidance more effective than CT guidance? *Eur J Radiol, 83*, 111-116.

Cohen, S. M., Shulman, R. G., & McLaughlin, A. C. (1979). Effects of ethanol on alanine metabolism in perfused mouse liver studied by 13C NMR. *Proceedings of the National Academy of Sciences, 76*, 4808-4812.

Colombo, C., Miceli, R., Lazar, A. J., Perrone, F., Pollock, R. E., Le Cesne, A., Hartgrink, H. H., Cleton-Jansen, A. M., Domont, J., Bovee, J. V., Bonvalot, S., Lev, D., & Gronchi, A. (2013). CTNNB1 45F mutation is a molecular prognosticator of increased postoperative primary desmoid tumor recurrence: an independent, multicenter validation study. *Cancer, 119*, 3696-3702.

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol, 26*, 1367-1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., & Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res, 10*, 1794-1805.

Deepali, J. (2015). Liver and intrahepatic bile ducts - tumor. In *Hepatocellular carcinoma - General* (Vol. 2017).

Degli Esposti, D., Hamelin, J., Bosselut, N., Saffroy, R., Sebagh, M., Pommier, A., Martel, C., & Lemoine, A. (2012). Mitochondrial roles and cytoprotection in chronic liver injury. *Biochem Res Int, 2012*, 387626.

Dowman, J. K., Tomlinson, J. W., & Newsome, P. N. (2010). Pathogenesis of non-alcoholic fatty liver disease. *QJM: An International Journal of Medicine, 103*, 71-83.

Doyen, J., Duranton-Tanneur, V., Hostein, I., Karanian-Philippe, M., Chevreau, C., Breibach, F., Coutts, M., Dadone, B., Saint-Paul, M. C., Gugenheim, J., Duffaud, F., & Pedeutour, F. (2016). Spatio-temporal genetic heterogeneity of CTNNB1 mutations in sporadic desmoid type fibromatosis lesions. *Virchows Arch, 468*, 369-374.

du Prel, J.-B., Röhrig, B., Hommel, G., & Blettner, M. (2010). Choosing Statistical Tests: Part 12 of a Series on Evaluation of Scientific Publications. *Deutsches Ärzteblatt International, 107*, 343-348.

Dubois, N., Bennoun, M., Allemand, I., Molina, T., Grimber, G., Daudet-Monsac, M., Abelanet, R., & Briand, P. (1991). Time-course development of differentiated hepatocarcinoma and lung metastasis in transgenic mice. *J Hepatol, 13*, 227-239.

Easton, D. F., Pharoah, P. D. P., Antoniou, A. C., Tischkowitz, M., Tavtigian, S. V., Nathanson, K. L., Devilee, P., Meindl, A., Couch, F. J., Southey, M., Goldgar, D. E., Evans, D. G. R., Chenevix-Trench, G., Rahman, N., Robson, M., Domchek, S. M., & Foulkes, W. D. (2015). Gene-Panel Sequencing and the Prediction of Breast-Cancer Risk. *New England Journal of Medicine, 372*, 2243-2257.

El-Serag, H. B. (2012). Epidemiology of Viral Hepatitis and Hepatocellular Carcinoma. *Gastroenterology, 142*, 1264-1273.e1261.

Elliott, W. H., & Elliott, D. C. (2005). *Biochemistry and molecular biology*. Oxford: Oxford University Press.

Feigelson, J., Pecau, Y., Cathelineau, L., & Navarro, J. (1975). Additional data on hepatic function tests in cystic fibrosis. *Acta Paediatr Scand, 64*, 337-344.

Ferrin, G., Rodriguez-Peralvarez, M., Aguilar-Melero, P., Ranchal, I., Llamoza, C., Linares, C. I., Gonzalez-Rubio, S., Muntane, J., Briceno, J., Lopez-Cillero, P., Montero-Alvarez, J. L., & de la Mata, M. (2015). Plasma protein biomarkers of hepatocellular carcinoma in HCV-infected alcoholic patients with cirrhosis. *PLoS One, 10*, e0118527.

Forner, A., Llovet, J. M., & Bruix, J. (2012). Hepatocellular carcinoma. *The Lancet, 379*, 1245-1255.

Fritz, A., Percy, C., Jack, A., Shanmuragatnam, K., Sobin, L. H., Parkin, D. M., & Whelan, S. (2000). *International Classification of Diseases for Oncology*. Geneva: World Health Organization.

Fujimoto, A., Totoki, Y., Abe, T., Boroevich, K. A., Hosoda, F., Nguyen, H. H., Aoki, M., Hosono, N., Kubo, M., Miya, F., Arai, Y., Takahashi, H., Shirakihara, T., Nagasaki, M., Shibuya, T., Nakano, K., Watanabe-Makino, K., Tanaka, H., Nakamura, H., Kusuda, J., Ojima, H., Shimada, K., Okusaka, T., Ueno, M., Shigekawa, Y., Kawakami, Y., Arihiro, K., Ohdan, H., Gotoh, K., Ishikawa, O., Ariizumi, S.-i., Yamamoto, M., Yamada, T., Chayama, K., Kosuge, T., Yamaue, H., Kamatani, N., Miyano, S., Nakagama, H., Nakamura, Y., Tsunoda, T., Shibata, T., & Nakagawa, H. (2012). Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat Genet, 44*, 760.

Gamallo, C., Palacios, J., Moreno, G., Calvo de Mora, J., Suarez, A., & Armas, A. (1999). beta-catenin expression pattern in stage I and II ovarian carcinomas : relationship with beta-catenin gene mutations, clinicopathological features, and clinical outcome. *Am J Pathol, 155*, 527-536.

Giannakis, M., Hodis, E., Jasmine Mu, X., Yamauchi, M., Rosenbluh, J., Cibulskis, K., Saksena, G., Lawrence, M. S., Qian, Z. R., Nishihara, R., Van Allen, E. M., Hahn, W. C., Gabriel, S. B., Lander, E. S., Getz, G., Ogino, S., Fuchs, C. S., & Garraway, L. A. (2014). RNF43 is frequently mutated in colorectal and endometrial cancers. *Nat Genet, 46*, 1264-1266.

GOconsortium, G. O. C. (2015). Gene Ontology Consortium: going forward. *Nucleic Acids Res, 43*, D1049-D1056.

Gough, D. A., Armour, J. C., & Baker, D. A. (1997). Advances and prospects in glucose assay technology. *Diabetologia, 40*, S102-S107.

Grant, D. M. (1991). Detoxification pathways in the liver. *J Inherit Metab Dis, 14*, 421-430.

Greten, T. F., Papendorf, F., Bleck, J. S., Kirchhoff, T., Wohlberedt, T., Kubicka, S., Klempnauer, J., Galanski, M., & Manns, M. P. (2005). Survival rate in patients with hepatocellular carcinoma: a retrospective analysis of 389 patients. *British Journal of Cancer, 92*, 1862-1868.

Guo, T., Kouvonen, P., Koh, C. C., Gillet, L. C., Wolski, W. E., Rost, H. L., Rosenberger, G., Collins, B. C., Blum, L. C., Gillessen, S., Joerger, M., Jochum, W., & Aebersold, R. (2015). Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. *21*, 407-413.

Gupta, S., Bent, S., & Kohlwes, J. (2003). Test characteristics of alpha-fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C. A systematic review and critical analysis. *Ann Intern Med, 139*, 46-50.

Hilgard, P., Muller, S., Hamami, M., Sauerwein, W. S., Haberkorn, U., Gerken, G., & Antoch, G. (2009). Selective internal radiotherapy (radioembolization) and radiation therapy for HCC-current status and perspectives. *Z Gastroenterol, 47*, 37-54.

Hirata, H., Sugimachi, K., Komatsu, H., Ueda, M., Masuda, T., Uchi, R., Sakimura, S., Nambara, S., Saito, T., Shinden, Y., Iguchi, T., Eguchi, H., Ito, S., Terashima, K., Sakamoto, K., Hirakawa, M., Honda, H., & Mimori, K. (2016). Decreased Expression of Fructose-1,6-bisphosphatase Associates with Glucose Metabolism and Tumor Progression in Hepatocellular Carcinoma. *Cancer Res, 76*, 3265-3276.

Huang, H., Liang, P., Yu, X. L., Cheng, Z. G., Han, Z. Y., Yu, J., & Liu, F. Y. (2015). Safety assessment and therapeutic efficacy of percutaneous microwave ablation therapy combined with

percutaneous ethanol injection for hepatocellular carcinoma adjacent to the gallbladder. *Int J Hyperthermia, 31*, 40-47.

IARC, I. A. f. R. o. C. (2012). GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. In *Cancer Fact Sheets*: International Agency for Research on Cancer.

Jacobs, J. M., Diamond, D. L., Chan, E. Y., Gritsenko, M. A., Qian, W., Stastna, M., Baas, T., Camp, D. G., 2nd, Carithers, R. L., Jr., Smith, R. D., & Katze, M. G. (2005). Proteome analysis of liver cells expressing a full-length hepatitis C virus (HCV) replicon and biopsy specimens of posttransplantation liver from HCV-infected patients. *J Virol, 79*, 7558-7569.

Janku, F., Wheler, J. J., Naing, A., Stepanek, V. M., Falchook, G. S., Fu, S., Garrido-Laguna, I., Tsimberidou, A. M., Piha-Paul, S. A., Moulder, S. L., Lee, J. J., Luthra, R., Hong, D. S., & Kurzrock, R. (2012). PIK3CA mutations in advanced cancers: characteristics and outcomes. *Oncotarget, 3*, 1566-1575.

Jelic, S., & Sotiropoulos, G. C. (2010). Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol, 21 Suppl 5*, v59-64.

Kammermeier, J., Drury, S., James, C. T., Dziubak, R., Ocaka, L., Elawad, M., Beales, P., Lench, N., Uhlig, H. H., Bacchelli, C., & Shah, N. (2014). Targeted gene panel sequencing in children with very early onset inflammatory bowel disease—evaluation and prospective analysis. *Journal of Medical Genetics, 51*, 748-755.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res, 45*, D353-d361.

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res, 28*, 27-30.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res, 44*, D457-462.

Keating, G. M. (2017). Sorafenib: A Review in Hepatocellular Carcinoma. *Target Oncol, 12*, 243-253.

Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik, 26*, 231-243.

Klotz, T., Montoriol, P. F., Da Ines, D., Petitcolin, V., Joubert-Zakeyh, J., & Garcier, J. M. (2013). Hepatic haemangioma: Common and uncommon imaging features. *Diagnostic and Interventional Imaging, 94*, 849-859.

Kobayashi, K., Sagae, S., Nishioka, Y., Tokino, T., & Kudo, R. (1999). Mutations of the beta-catenin gene in endometrial carcinomas. *Jpn J Cancer Res, 90*, 55-59.

Kolble, K., Barthel, B., Ullrich, O., Pidde, H., Dohring, C., Ruschoff, J., Schlag, P. M., & Dietel, M. (2000). beta-Catenine as a genomic target of high-grade microsatellite instability in colorectal cancer. *Verh Dtsch Ges Pathol, 84*, 182-186.

Kuich, P. H., Hoffmann, N., & Kempa, S. (2014). Maui-VIA: A User-Friendly Software for Visual Identification, Alignment, Correction, and Quantification of Gas Chromatography-Mass Spectrometry Data. *Front Bioeng Biotechnol, 2*, 84.

Lambert, B., & Van De Wiele, C. (2009). Selective internal radiation therapy of HCC and liver metastases: a locoregional or worldwide therapy? *Q J Nucl Med Mol Imaging, 53*, 302-304.

Lee, J. G., Kang, C. M., Park, J. S., Kim, K. S., Yoon, D. S., Choi, J. S., Lee, W. J., & Kim, B. R. (2006). The Actual Five-year Survival Rate of Hepatocellular Carcinoma Patients after Curative Resection. *Yonsei Medical Journal, 47*, 105-112.

Lin, C. Y., Wu, H., Tjeerdema, R. S., & Viant, M. R. (2007). Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. *Metabolomics, 3*, 55-67.

Liu, Y., & Wu, F. (2010). Global burden of aflatoxin-induced hepatocellular carcinoma: a risk assessment. *Environ Health Perspect, 118*, 818-824.

Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J. F., de Oliveira, A. C., Santoro, A., Raoul, J. L., Forner, A., Schwartz, M., Porta, C., Zeuzem, S., Bolondi, L., Greten, T. F., Galle, P. R., Seitz, J. F., Borbath, I., Haussinger, D., Giannaris, T., Shan, M., Moscovici, M., Voliotis, D., & Bruix, J. (2008). Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med, 359*, 378-390.

Lobo, L., Yakoub, D., Picado, O., Ripat, C., Pendola, F., Sharma, R., ElTawil, R., Kwon, D., Venkat, S., Portelance, L., & Yechieli, R. (2016). Unresectable Hepatocellular Carcinoma: Radioembolization Versus Chemoembolization: A Systematic Review and Meta-analysis. *Cardiovasc Intervent Radiol, 39*, 1580-1588.

Ludwig, J., Viggiano, T. R., McGill, D. B., & Oh, B. J. (1980). Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clin Proc, 55*, 434-438.

Luk, J. M., & Liu, A. M. (2011). Proteomics of hepatocellular carcinoma in Chinese patients. *Omics, 15*, 261-266.

Margini, C., & Dufour, J. F. (2016). The story of HCC in NAFLD: from epidemiology, across pathogenesis, to prevention and treatment. *Liver Int, 36*, 317-324.

Martínez-Granados, B., Monleón, D., Martínez-Bisbal, M. C., Rodrigo, J. M., Olmo, J. d., Lluch, P., Ferrández, A., Martí-Bonmatí, L., & Celda, B. (2006). Metabolite identification in human liver needle biopsies by high-resolution magic angle spinning 1H NMR spectroscopy. *NMR in Biomedicine, 19*, 90-100.

Maruyama, Y., Okuda, K., Ogata, T., Yasunaga, M., Ishikawa, H., Hirakawa, Y., Fukuyo, K., Horiuchi, H., Nakashima, O., & Kinoshita, H. (2013). Perioperative Challenges and Surgical Treatment of Large Simple, and Infectious Liver Cyst - A 12-Year Experience. *PLoS One, 8*, e76537.

Mastrobuoni, G., Irgang, S., Pietzke, M., Assmus, H. E., Wenzel, M., Schulze, W. X., & Kempa, S. (2012). Proteome dynamics and early salt stress response of the photosynthetic organism Chlamydomonas reinhardtii. *BMC Genomics, 13*, 215.

Masuda, T., & Miyoshi, E. (2011). Cancer biomarkers for hepatocellular carcinomas: from traditional markers to recent topics. *Clin Chem Lab Med, 49*, 959-966.

MeSH, U. S. N. L. o. M. (1999). Medical Subject Headings (MeSH®). In (01.01.1999 ed.). 8600 Rockville Pike, Bethesda, MD 20894: U.S. National Library of Medicine.

Mondal, G., Saroha, A., Bose, P. P., & Chatterjee, B. P. (2016). Altered glycosylation, expression of serum haptoglobin and alpha-1-antitrypsin in chronic hepatitis C, hepatitis C induced liver cirrhosis and hepatocellular carcinoma patients. *Glycoconj J, 33*, 209-218.

Moreno-Bueno, G., Gamallo, C., Perez-Gallego, L., de Mora, J. C., Suarez, A., & Palacios, J. (2001). beta-Catenin expression pattern, beta-catenin gene mutations, and microsatellite instability in endometrioid ovarian carcinomas and synchronous endometrial carcinomas. *Diagn Mol Pathol, 10*, 116-122.

Mouradov, D., Sloggett, C., Jorissen, R. N., Love, C. G., Li, S., Burgess, A. W., Arango, D., Strausberg, R. L., Buchanan, D., Wormald, S., O'Connor, L., Wilding, J. L., Bicknell, D., Tomlinson, I. P., Bodmer, W. F., Mariadason, J. M., & Sieber, O. M. (2014). Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. *Cancer Res, 74*, 3238-3247.

Mullen, J. T., DeLaney, T. F., Rosenberg, A. E., Le, L., Iafrate, A. J., Kobayashi, W., Szymonifka, J., Yeap, B. Y., Chen, Y. L., Harmon, D. C., Choy, E., Yoon, S. S., Raskin, K. A., Hornicek, F. J., & Nielsen, G. P. (2013). beta-Catenin mutation status and outcomes in sporadic desmoid tumors. *Oncologist, 18*, 1043-1049.

Nelson, D. L., Nelson, D. L., Lehninger, A. L., & Cox, M. M. (2008). *Lehninger principles of biochemistry*. New York: W.H. Freeman.

Ng, P. C., & Kirkness, E. F. (2010). Whole Genome Sequencing. In M. R. Barnes & G. Breen (Eds.), *Genetic Variation: Methods and Protocols* (pp. 215-226). Totowa, NJ: Humana Press.

Nguyen, V. T., Law, M. G., & Dore, G. J. (2009). Hepatitis B-related hepatocellular carcinoma: epidemiological characteristics and disease burden. *J Viral Hepat, 16*, 453-463.

Nishida, Y., Tsukushi, S., Urakawa, H., Hamada, S., Kozawa, E., Ikuta, K., & Ishiguro, N. (2016). Simple resection of truncal desmoid tumors: A case series. *Oncol Lett, 12*, 1564-1568.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem, 250*, 4007-4021.

Orimo, T., Ojima, H., Hiraoka, N., Saito, S., Kosuge, T., Kakisaka, T., Yokoo, H., Nakanishi, K., Kamiyama, T., Todo, S., Hirohashi, S., & Kondo, T. (2008). Proteomic profiling reveals the prognostic value of adenomatous polyposis coli-end-binding protein 1 in hepatocellular carcinoma. *Hepatology, 48*, 1851-1863.

Ouda, S. M., Khairy, A. M., Sorour, A. E., & Mikhail, M. N. (2015). Serum Beta-2 Microglobulin: a Possible Marker for Disease Progression in Egyptian Patients with Chronic HCV Related Liver Diseases. *Asian Pac J Cancer Prev, 16*, 7825-7829.

Palacios, J., & Gamallo, C. (1998). Mutations in the beta-catenin gene (CTNNB1) in endometrioid ovarian carcinomas. *Cancer Res, 58*, 1344-1347.

Pandya, J. D., Sullivan, P. G., Leung, L. Y., Tortella, F. C., Shear, D. A., & Deng-Bryant, Y. (2016). Advanced and High-Throughput Method for Mitochondrial Bioenergetics Evaluation in Neurotrauma. In F. H. Kobeissy, C. E. Dixon, R. L. Hayes & S. Mondello (Eds.), *Injury Models of the Central Nervous System: Methods and Protocols* (pp. 597-610). New York, NY: Springer New York.

Paoletti, A., & Chang, F. (2000). Analysis of mid1p, a Protein Required for Placement of the Cell Division Site, Reveals a Link between the Nucleus and the Cell Surface in Fission Yeast. *Molecular Biology of the Cell, 11*, 2757-2773.

Paul, S. B., Gulati, M. S., Sreenivas, V., Madan, K., Gupta, A. K., Mukhopadhyay, S., & Acharya, S. K. (2007). Evaluating patients with cirrhosis for hepatocellular carcinoma: value of clinical symptomatology, imaging and alpha-fetoprotein. *Oncology, 72 Suppl 1*, 117-123.

Pei, Y., Zhang, T., Renault, V., & Zhang, X. (2009). An overview of hepatocellular carcinoma study by omics-based methods. *Acta Biochim Biophys Sin (Shanghai), 41*, 1-15.

Peterson, K. (2004). Biomarkers for alcohol use and abuse--a summary. *Alcohol Res Health, 28*, 30-37.

Pietzke, M., Zasada, C., Mudrich, S., & Kempa, S. (2014). Decoding the dynamics of cellular metabolism and the action of 3-bromopyruvate and 2-deoxyglucose using pulsed stable isotope-resolved metabolomics. *Cancer Metab, 2*, 9.

Powell, E. E., Cooksley, W. G., Hanson, R., Searle, J., Halliday, J. W., & Powell, L. W. (1990). The natural history of nonalcoholic steatohepatitis: a follow-up study of forty-two patients for up to 21 years. *Hepatology, 11*, 74-80.

QIAGEN. (2006). *DNeasy® Blood & Tissue Handbook*: QIAGEN.

Rappsilber, J., Ishihama, Y., & Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem, 75*, 663-670.

Reisch, A. S., & Elpeleg, O. (2007). Biochemical Assays for Mitochondrial Activity: Assays of TCA Cycle Enzymes and PDHc. In *Methods in Cell Biology* (Vol. 80, pp. 199-222): Academic Press.

Riener, M. O., Bawohl, M., Clavien, P. A., & Jochum, W. (2008). Rare PIK3CA hotspot mutations in carcinomas of the biliary tract. *Genes Chromosomes Cancer, 47*, 363-367.

Roytberg, G. E., & Strutynskiy, A. V. (2013). *Liver, bile ducts, pancreas. Textbook.* Moscow, Russian Federation: МЕДпресс-информ.

Shang, R. Z., Qu, S. B., & Wang, D. S. (2016). Reprogramming of glucose metabolism in hepatocellular carcinoma: Progress and prospects. *World J Gastroenterol, 22*, 9933-9943.

Shang, S., Qin, X., Li, W., Zhang, S., & Liu, Y. (2016). ELISA index of serum fucosylated haptoglobin for diagnosis of HCC using the normal and reverse AAL ELISA. *Discov Med, 21*, 15-23.

Shen, A., Tang, C., Wang, Y., Chen, Y., Yan, X., Zhang, C., Liu, R., Wei, X., Zhu, Y., Zhang, H., & Wu, Z. (2013). A systematic review of sorafenib in Child-Pugh A patients with unresectable hepatocellular carcinoma. *J Clin Gastroenterol, 47*, 871-880.

Shen, J. Y., Li, C., Wen, T. F., Yan, L. N., Li, B., Wang, W. T., Yang, J. Y., Xu, M. Q., & Nazar Highness, T. (2016). Liver transplantation versus surgical resection for HCC meeting the Milan criteria: A propensity score analysis. *Medicine (Baltimore), 95*, e5756.

Snoep, J. L., & Westerhoff, H. V. (2005). From isolation to integration, a systems biology approach for building the Silicon Cell. In L. Alberghina & H. V. Westerhoff (Eds.), *Systems Biology: Definitions and Perspectives* (pp. 13-30). Berlin, Heidelberg: Springer Berlin Heidelberg.

Sudarsan, S., Dethlefsen, S., Blank, L. M., Siemann-Herzberg, M., & Schmid, A. (2014). The Functional Structure of Central Carbon Metabolism in Pseudomonas putida KT2440. *Applied and Environmental Microbiology, 80*, 5292-5303.

Tan, G. S., Lim, K. H., Tan, H. T., Khoo, M. L., Tan, S. H., Toh, H. C., & Ching Ming Chung, M. (2014). Novel proteomic biomarker panel for prediction of aggressive metastatic hepatocellular carcinoma relapse in surgically resectable patients. *J Proteome Res, 13*, 4833-4846.

Tawara, S., Tatsumi, T., Iio, S., Kobayashi, I., Shigekawa, M., Hikita, H., Sakamori, R., Hiramatsu, N., Miyoshi, E., & Takehara, T. (2016). Evaluation of Fucosylated Haptoglobin and Mac-2 Binding Protein as Serum Biomarkers to Estimate Liver Fibrosis in Patients with Chronic Hepatitis C. *PLoS One, 11*, e0151828.

The Genomes Project, C. (2015). A global reference for human genetic variation. *Nature, 526*, 68-74.

Tian, Q., Price, N. D., & Hood, L. (2012). Systems cancer medicine: towards realization of predictive, preventive, personalized and participatory (P4) medicine. *J Intern Med, 271*, 111-121.

Tietze, F. (1969). Enzymic method for quantitative determination of nanogram amounts of total and oxidized glutathione: Applications to mammalian blood and other tissues. *Anal Biochem, 27*, 502-522.

Tolosa, L., Malak, H., Raob, G., & Lakowicz, J. R. (1997). Optical assay for glucose based on the luminescnence decay time of the long wavelength dye Cy5™. *Sensors and Actuators B: Chemical, 45*, 93-99.

Tornesello, M. L., Buonaguro, L., Tatangelo, F., Botti, G., Izzo, F., & Buonaguro, F. M. (2013). Mutations in TP53, CTNNB1 and PIK3CA genes in hepatocellular carcinoma associated with hepatitis B and hepatitis C virus infections. *Genomics, 102*, 74-83.

Ueno, M., Hayami, S., Shigekawa, Y., Kawai, M., Hirono, S., Okada, K., Tamai, H., Shingaki, N., Mori, Y., Ichinose, M., & Yamaue, H. (2015). Prognostic impact of surgery and radiofrequency ablation on single nodular HCC 5 cm: Cohort study based on serum HCC markers. *J Hepatol, 63*, 1352-1359.

Ukawa, S., Okada, E., Nakamura, K., Hirata, M., Nagai, A., Matsuda, K., Yamagata, Z., Kamatani, Y., Ninomiya, T., Kiyohara, Y., Muto, K., Kubo, M., Nakamura, Y., & Tamakoshi, A. (2017). Characteristics of patients with liver cancer in the BioBank Japan project. *J Epidemiol, 27*, S43-s48.

UniProtconsortium. (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res, 45*, D158-D169.

Vander Heiden, M. G., Cantley, L. C., & Thompson, C. B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science, 324*, 1029-1033.

VIB/UGent, B. E. G. (2017). Draw Venn Diagram. In: VIB/UGent Bioinformatics & Evolutionary Genomics.

Vucenik, I., & Stains, J. P. (2012). Obesity and cancer risk: evidence, mechanisms, and recommendations. *Annals of the New York Academy of Sciences, 1271*, 37-43.

Warburg, O. (1956). On the origin of cancer cells. *Science, 123*, 309-314.

Willekens, I., Hoorens, A., Geers, C., Op de Beeck, B., Vandenbroucke, F., & de Mey, J. (2009). Combined hepatocellular and cholangiocellular carcinoma presenting with radiological characteristics of focal nodular hyperplasia. *World Journal of Gastroenterology : WJG, 15*, 3940-3943.

Yang, B., Zan, R. Y., Wang, S. Y., Li, X. L., Wei, M. L., Guo, W. H., You, X., Li, J., & Liao, Z. Y. (2015). Radiofrequency ablation versus percutaneous ethanol injection for hepatocellular carcinoma: a meta-analysis of randomized controlled trials. *World J Surg Oncol, 13*, 96.

Yizhak, K., Benyamini, T., Liebermeister, W., Ruppin, E., & Shlomi, T. (2010). Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. *Bioinformatics, 26*, i255-i260.

Yu, D., Shi, X., Meng, G., Chen, J., Yan, C., Jiang, Y., Wei, J., & Ding, Y. (2015). Kidney-type glutaminase (GLS1) is a biomarker for pathologic diagnosis and prognosis of hepatocellular carcinoma. *Oncotarget, 6*, 7619-7631.

Zinkin, N. T., Grall, F., Bhaskar, K., Otu, H. H., Spentzos, D., Kalmowitz, B., Wells, M., Guerrero, M., Asara, J. M., Libermann, T. A., & Afdhal, N. H. (2008). Serum proteomics and biomarkers in hepatocellular carcinoma and chronic liver disease. *Clin Cancer Res, 14*, 470-477.

Contributions of collaborators

The mice material for proteomic and metabolomics analyses was provided by our Prof. Dr. Thorsten Cramer and our colleagues from his laboratory, Charite Medical University, Berlin, Germany.

Human samples were obtained in collaboration with Prof. Dr. Thorsten Cramer and his colleagues from University Hospital Aachen.

The work with the NGS was performed in the lab of Prof. Dr. Christine Sers, Charite Medical University, Berlin, Germany. NGS runs were performed with the help of Andrea Menne, and data was analyzed with the help of Dr. Soulafa Mamlouk.

List of publications

- Kettelhake A, Berndt N, Vvedenskaya O, Cramer T: Application of mathematical modelling to assess the functional importance of HIF1A for hepatoma biology. *Z Gastroenterol*, 2015; 53 - A4_7;
- Samhan-Arias AK, Ji J, Demidova O, Kagan VE, Amoscato AA.: Oxidized phospholipids as biomarkers of tissue and cell damage with a focus on cardiolipin. *Biochem. Biophys. Acta*, 2012, Oct; 1818 (10):2413-23;
- Demidova O, Stepanov GO and Osipov AN: How can free radicals activate apoptotic reaction? *Bulletin of RSMU*, 2010, 2:433.

Selected conferences and schools

- Hallmarks of Cancer, Ghent, Belgium, 2016, “Metabolic reprogramming of hepatocellular carcinoma” (poster);
- Young Scientists in Cancer Symposium, Berlin, Germany, 2015, “Metabolic reprogramming of hepatocellular carcinoma” (talk);
- Hypoxia: from basic mechanisms to therapeutics, Royal Dublin Society, Dublin, Ireland, 2015, “Application of mathematical modelling to assess the functional importance of HIF-1a for hepatoma biology” (poster coauthor);
- Quantitative Proteomics and Data Analysis workshop, Chester, UK, 2014 (participant);
- The SignGene Winter School "Biology of Differentiation and Cancer" Haifa, Israel, 2014, “Different characteristics of metabolism of Hepatocellular Carcinoma cell lines” (poster);
- Society of Toxicology’s 51st Annual Meeting & ToxExpo, San Francisco, California, USA, 2012, “Normal-phase/reverse-phase 2D-HPLC-MS analysis of oxidized lipid species: Application to the study of cardiolipin in apoptosis and damaged tissue” (poster coauthor).

Appendix 1

HCC macroscopic picture and histology

Macroscopic picture

Figure 1. Hepatocellular carcinoma with a greenish yellow hue (Deepali, 2015).

Histological features (Deepali, 2015)

- Patterns are trabecular (most common) with 4 and more cells surrounded by layer of flattened endothelial cells; also clear cell, giant cell;
- Presence of sinusoidal vessels surrounding tumor cells;
- Scanty (poor) stroma;
- Cells are polygonal with distinct cell membranes, abundant granular eosinophilic cytoplasm, round nuclei with rough chromatin and thickened nuclear membrane;
- Presence of portal vein thrombosis, vascular invasion;
- Variable: abundant fat, bile (5-33%);
- Minimal desmoplasia.

Histological types (Deepali, 2015)

- **Well differentiated:** thin plates (1-3 hepatocytes thick), cells smaller than normal, minimal nuclear atypia, nuclear density 2x normal liver; commonly change of liver composition; hepatocyte adenoma may be present; common pattern for small hepatocellular carcinoma;
- **Moderately differentiated:** trabecular pattern with 4 and more cells thick; larger tumor cells than well-differentiated HCC with more eosinophilic cytoplasm, distinct nucleoli, bile and tumor giant cells; most common pattern in advanced HCC;
- **Poorly differentiated:** large tumor cells with hyperchromatic nuclei in compact growth pattern with rare trabeculae or bile; may not seem to be hepatocellular;
- **Combined hepatocellular-cholangiocarcinoma:** <1% of primary liver carcinomas; unequivocal hepatocellular and cholangiocarcinoma that are intimately admixed(Willekens, et al., 2009); increased CA19-9 and AFP;
- **Diffuse cirrhosis like HCC:** diffuse and extensive liver involvement by small cirrhosis-like nodules that evade clinical and radiographic detection (Jakate, et al., 2010);
- **Steatohepatitic HCC:** recently described variant associated with metabolic dysfunction such as non-alcoholic fatty liver disease (Salomao, et al., 2012; Salomao, Yu, Brown, Emond, & Lefkowitch, 2010); slightly firmer than classic HCC (due to fibrosis) and more yellow (due to steatosis); histology shows steatosis, hepatocyte ballooning, inflammation and pericellular fibrosis within neoplastic tissue.

Examples of histological pictures

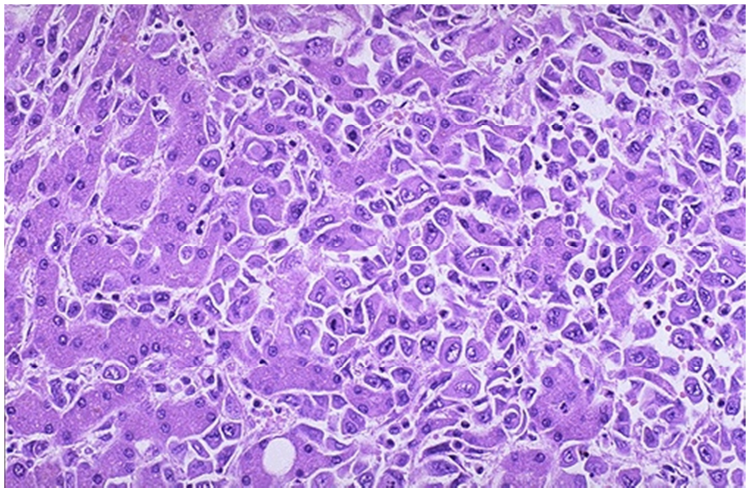


Figure 2. The malignant cells of this hepatocellular carcinoma (seen mostly on the right) are well differentiated and interdigitate with normal, larger hepatocytes (seen mostly at the left) (Deepali, 2015).

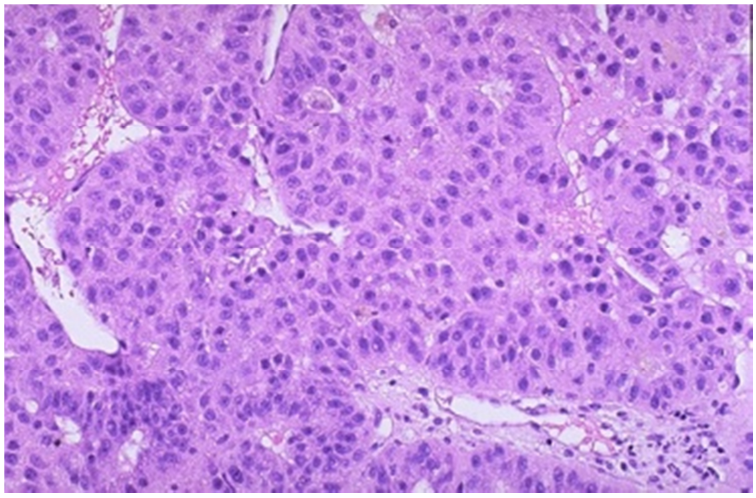


Figure 3. Hepatocellular carcinoma, histological picture. Hepatocellular carcinoma is presented with liver cords that are wider than the normal liver plate that is two cells thick. No normal lobular architecture is observed on this picture, despite presence of vascular structures (Deepali, 2015).

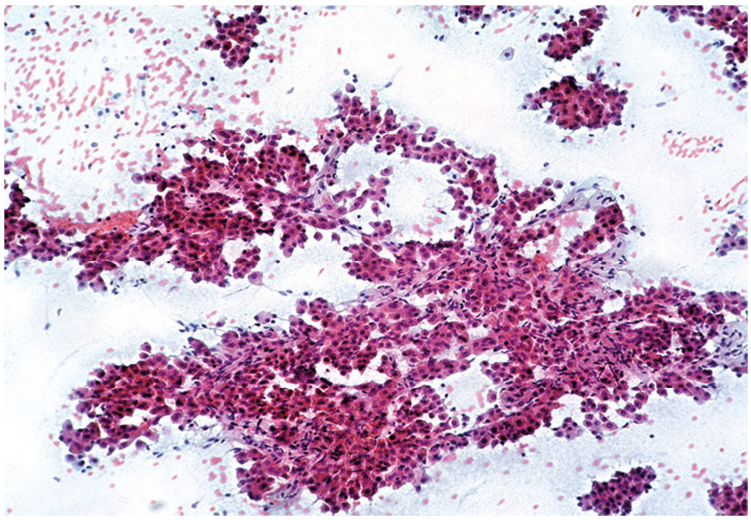


Figure 4. Hepatocellular carcinoma, histological picture. The cohesive clusters of malignant hepatocytes appear as slender arborizing cords (Bosman, et al., 2010).

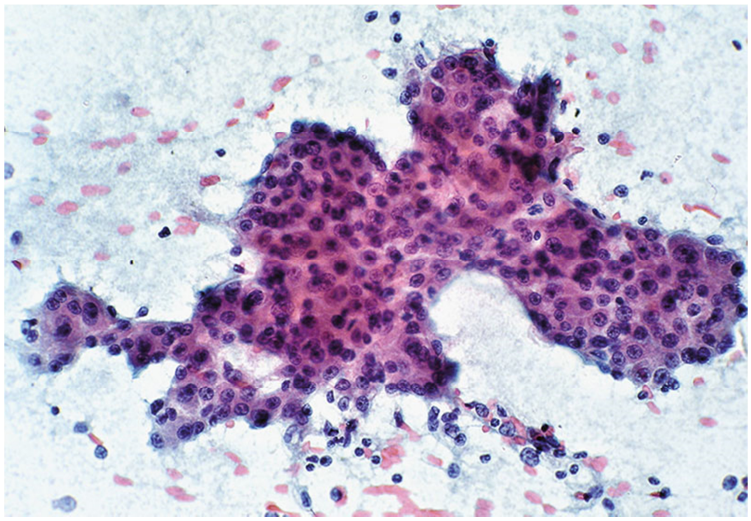


Figure 5. Hepatocellular carcinoma, histological picture. The clusters of malignant hepatocytes with broad cords (more than five cells thick) are wrapped by endothelium. Note the granular cytoplasm, increased nucleus/cytoplasm ratio, granular chromatin and distinct nucleolus (Bosman, et al., 2010).

Bibliography

Bosman, F. T., Carneiro, F., Hruban, R. H., & Theise, N. D. (2010). *WHO Classification of Tumours of the Digestive System* (4 ed. Vol. 3).

Deepali, J. (2015). Liver and intrahepatic bile ducts - tumor. In *Hepatocellular carcinoma - General* (Vol. 2017).

Jakate, S., Yabes, A., Giusto, D., Naini, B., Lassman, C., Yeh, M. M., & Ferrell, L. D. (2010). Diffuse cirrhosis-like hepatocellular carcinoma: a clinically and radiographically undetected variant mimicking cirrhosis. *Am J Surg Pathol, 34*, 935-941.

Salomao, M., Remotti, H., Vaughan, R., Siegel, A. B., Lefkowitch, J. H., & Moreira, R. K. (2012). The steatohepatitic variant of hepatocellular carcinoma and its association with underlying steatohepatitis. *Hum Pathol, 43*, 737-746.

Salomao, M., Yu, W. M., Brown, R. S., Jr., Emond, J. C., & Lefkowitch, J. H. (2010). Steatohepatitic hepatocellular carcinoma (SH-HCC): a distinctive histological variant of HCC in hepatitis C virus-related cirrhosis with associated NAFLD/NASH. *Am J Surg Pathol, 34*, 1630-1636.

Willekens, I., Hoorens, A., Geers, C., Op de Beeck, B., Vandenbroucke, F., & de Mey, J. (2009). Combined hepatocellular and cholangiocellular carcinoma presenting with radiological characteristics of focal nodular hyperplasia. *World Journal of Gastroenterology : WJG, 15*, 3940-3943.

Appendix 2

Table. Full names and short names of chosen enzymes of central carbon metabolism.

Protein names	**Gene names (*homo sapiens*)**	**Protein IDs (*homo sapiens*)**	**Gene names (*mus musculus*)**	**Protein IDs (*mus musculus*)**
ATP-citrate synthase	ACLY	P53396;P53396-2	Acly	Q91V92
Cytoplasmic aconitate hydratase	ACO1	P21399	Aco1	P28271
Aconitate hydratase, mitochondrial	ACO2	Q99798	Aco2	Q99KI0
Long-chain-fatty-acid--CoA ligase 1	ACSL1	P33121;P33121-2	Acsl1	P41216
Long-chain-fatty-acid--CoA ligase 3	ACSL3	O95573	Acsl3	Q9CZW4
Long-chain-fatty-acid--CoA ligase 4	ACSL4	O60488-2;O60488	Acsl4	Q9QUJ7-2;Q9QUJ7
Long-chain-fatty-acid--CoA ligase 5	ACSL5	Q9ULC5;Q9ULC5-3	Acsl5	Q8JZR0
Long-chain-fatty-acid--CoA ligase 6	ACSL6	Q9UKU0;Q9UKU0-(1,2,3,5,6,7,8)	Acsl6	Q91WC3
Fructose-bisphosphate aldolase A	ALDOA	P04075	Aldoa	P05064
Fructose-bisphosphate aldolase B	ALDOB	P05062	Aldob	Q91Y97
Fructose-bisphosphate aldolase C	ALDOC	P09972	Aldoc	P05063
Citrate synthase, mitochondrial	CS	O75390	Cs	Q9CZU6
Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	DLAT	P10515	Dlat	Q8BMF4
Dihydrolipoyl dehydrogenase, mitochondrial	DLD	P09622	Dld	O08749
Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	DLST	P36957	Dlst	Q9D2G2;Q9D2G2-2
Alpha-enolase	ENO1	P06733	Eno1	P17182
Gamma-enolase	ENO2	P09104	Eno2	P17183
Beta-enolase	ENO3	P13929;P13929-2;P13929-3	Eno3	P21550
Fatty acid synthase	FASN	P49327	Fasn	P19096
Fructose-1,6-bisphosphatase 1	FBP1	P09467	Fbp1	Q9QXD6
Fructose-1,6-bisphosphatase isozyme 2	FBP2	O00757	Fbp2	P70695
Fumarate hydratase, mitochondrial	FH	P07954-2;P07954	Fh	P97807-2;P97807
Glucose-6-phosphatase	G6PC	P35575	G6pc	P35576
Glucose-6-phosphate 1-dehydrogenase	G6PD	P11413;P11413-2;P11413-3	G6pdx	Q00612
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	P04406;O14556	Gapdh	P16858
Glucokinase	GCK	P35557-2;P35557-3	Gck	P52792-2;P52792
Glutaminase kidney isoform, mitochondrial	GLS	O94925	Gls	D3Z7P3;D3Z7P3-2
Glutaminase liver isoform, mitochondrial	GLS2	Q9UI32	Gls2	Q571F8
Glutamate dehydrogenase 1, mitochondrial	GLUD1	P00367	Glud1	P26443
Glutamate dehydrogenase 2, mitochondrial	GLUD2	P49448	-	
Glucose-6-phosphate isomerase	GPI	P06744;P06744-2	Gpi	P06745
Alanine aminotransferase 1	GPT	P24298	Gpt	Q8QZR5
Alanine aminotransferase 2	GPT2	Q8TD30;Q8TD30-2	Gpt2	Q8BGT5
Glycogen [starch] synthase, muscle	GYS1	P13807-2;P13807	Gys1	
Glycogen [starch] synthase, liver	GYS2	P54840	Gys2	Q8VCB3
Hexokinase-1	HK1	P19367;P19367-4;P19367-2;P19367-3	Hk1	P17710
Hexokinase-2	HK2	P52789	Hk2	O08528
Hexokinase-3	HK3	P52790	Hk3	Q3TRM8
Isocitrate dehydrogenase [NADP] cytoplasmic	IDH1	O75874	Idh1	O88844
Isocitrate dehydrogenase [NADP], mitochondrial	IDH2	P48735	Idh2	P54071
Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	IDH3A	P50213;P50213-2	Idh3a	Q9D6R2;Q9D6R2-2
Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial	IDH3B	O43837;O43837-2	Idh3b	
Isocitrate dehydrogenase [NAD] subunit gamma, mitochondrial	IDH3G	P51553	Idh3g	P70404
L-lactate dehydrogenase A chain	LDHA	P00338;P00338-3;P00338-4;P00338-5;P00338-2	Ldha	P06151
L-lactate dehydrogenase B chain	LDHB	P07195	Ldhb	P16125

L-lactate dehydrogenase C chain	LDHC	P07864	Ldhc	P00342
Probable D-lactate dehydrogenase, mitochondrial	LDHD	Q86WU2-2;Q86WU2	Ldhd	Q7TNG8
Malate dehydrogenase, cytoplasmic	MDH1	P40925;P40925-3;P40925-2	Mdh1	P14152
Malate dehydrogenase, mitochondrial	MDH2	P40926	Mdh2	P08249
NADP-dependent malic enzyme	ME1	P48163	Me1	P06801
NAD-dependent malic enzyme, mitochondrial	ME2	P23368;P23368-2	Me2	Q99KE1
NADP-dependent malic enzyme, mitochondrial	ME3	Q16798	Me3	Q8BMF3
Mitochondrial pyruvate carrier 1	MPC1	Q9Y5U8	Mpc1	P63030
Mitochondrial pyruvate carrier 2	MPC2	O95563	Mpc2	Q9D023
2-oxoglutarate dehydrogenase, mitochondrial	OGDH	Q02218;Q02218-2	Ogdh	Q60597;Q60597-3;Q60597-2; Q60597-4
2-oxoglutarate dehydrogenase-like, mitochondrial	OGDHL	Q9ULD0;Q9ULD0-2;Q9ULD0-3	Ogdhl	B2RXT3
Pyruvate carboxylase, mitochondrial	PC	P11498	Pc	Q05920
Phosphoenolpyruvate carboxykinase, cytosolic [GTP]	PCK1	P35558	Pck1	Q9Z2V4
Phosphoenolpyruvate carboxykinase [GTP], mitochondrial	PCK2	Q16822;Q16822-2	Pck2	Q8BH04
Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	PDHA1	P08559;P08559-3;P08559-2;P08559-4	Pdha1	P35486
Pyruvate dehydrogenase E1 component subunit alpha, testis-specific form, mitochondrial	PDHA2	P29803	Pdha2	P35487
Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	PDHB	P11177-2;P11177;P11177-3	Pdhb	Q9D051
Pyruvate dehydrogenase protein X component, mitochondrial	PDHX	O00330;O00330-2	Pdhx	Q8BKZ9
6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 1; 6-phosphofructo-2-kinase;Fructose-2,6-bisphosphatase	PFKFB1	P16118;Q16877;Q16875-2;Q16875	Pfkfb1	P70266;P70266-2
6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 2; 6-phosphofructo-2-kinase;Fructose-2,6-bisphosphatase	PFKFB2	O60825;O60825-2	Pfkfb2	P70265
ATP-dependent 6-phosphofructokinase, liver type	PFKL	P17858;P17858-2	Pfkl	P12382
ATP-dependent 6-phosphofructokinase, muscle type	PFKM	P08237;P08237-3;P08237-2	Pfkm	P47857
ATP-dependent 6-phosphofructokinase, platelet type	PFKP	Q01813;Q01813-2	Pfkp	Q9WUA3;Q9WUA3-2
Phosphoglycerate mutase 1	PGAM1	P18669	Pgam1	Q9DBJ1
Phosphoglycerate mutase 2	PGAM2	P15259	Pgam2	O70250
Serine/threonine-protein phosphatase PGAM5, mitochondrial	PGAM5	Q96HS1;Q96HS1-2	Pgam5	Q8BX10-2;Q8BX10
6-phosphogluconate dehydrogenase, decarboxylating	PGD	P52209	Pgd	Q9DCD0
Phosphoglycerate kinase 1	PGK1	P00558	Pgk1	P09411
Phosphoglycerate kinase 2	PGK2	P07205	Pgk2	P09041
Phosphoglucomutase-1	PGM1	P36871;P36871-2;P36871-3	Pgm1	Q9D0F9
Phosphoglucomutase-2	PGM2	Q96G03	Pgm2	Q7TSV4
Pyruvate kinase PKLR	PKLR	P30613	Pklr	P53657
Pyruvate kinase PKM	PKM	P14618;P14618-3	Pkm	P52480
Glycogen phosphorylase, liver form	PYGL	P06737-2;P06737	Pygl	Q9ET01
Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	SDHA	P31040	Sdha	Q8K2B3
Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	SDHB	P21912	Sdhb	Q9CQA3
Succinate dehydrogenase cytochrome b560 subunit, mitochondrial	SDHC	Q99643-5;Q99643-3;Q99643	Sdhc	Q9CZB0
Succinate dehydrogenase [ubiquinone] cytochrome b small subunit, mitochondrial	SDHD	O14521	Sdhd	Q9CXV1
Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial	SUCLA2	Q9P2R7-2;Q9P2R7	Sucla2	Q9Z2I9
Succinyl-CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial	SUCLG1	P53597	Suclg1	Q9WUM5
Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial	SUCLG2	Q96I99;Q96I99-2	Suclg2	Q9Z2I8;Q9Z2I8-2
Transaldolase	TALDO1	P37837	Taldo1	Q93092
Transketolase	TKT	P29401;P29401-2	Tkt	P40142
Transketolase-like protein 1	TKTL1	P51854;P51854-1	Tktl1	Q99MX0
Transketolase-like protein 2	TKTL2	Q9H0I9	Tktl2	Q9D4D4
Triosephosphate isomerase	TPI1	P60174-1;P60174;P60174-4	Tpi1	P17751

Appendix 3

Table. Patients list.

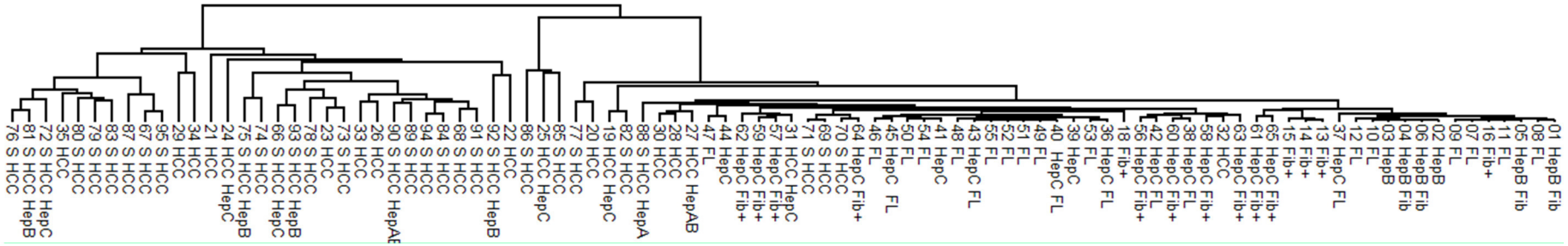
Sample Code	Acquisition Date	B/S	Diagnosis (short)	Diagnosis (full)	Mutations detected	Mutations confirmed	Age	Gender	Proteome	Metabolome	Genome
1	2-Nov-09	B	HBV, Fibrosis	Chronic Hepatitis B, Fibrosis	KIT		63	m	✓	✓	✓
2	18-Feb-09	B	HBV	Chronic Hepatitis B	RB1 KDR TP53	TP53 RB1	57	f	✓	✓	✓
3	25-Feb-09	B	HBV	Chronic Hepatitis B	TP53		44	m	✓	✓	✓
4	3-Apr-09	B	HBV, Fibrosis	Chronic Hepatitis B, mild Fibrosis	TP53		45	f	✓	✓	✓
5	29-Jul-09	B	HBV, Fibrosis	Chronic Hepatitis B, Fibrosis	TP53 KDR		45	m	✓	✓	✓
6	5-Nov-10	B	HBV, Fibrosis	Chronic Hepatitis B, Fibrosis	TP53		34	f	✓	✓	✓
7	13-Feb-08	B	Fatty Liver	Fatty Liver (70%)	TP53 PIK3CA KDR		47	f	✓		✓
8	26-Feb-08	B	Fatty Liver	Fatty Liver	TP53 MET		49	m	✓	✓	✓
9	3-Dec-08	B	Fatty Liver	Fatty Liver	STK11 KDR	STK11	59	f	✓	✓	✓
10	24-Jun-09	B	Fatty Liver	Fatty Liver	PIK3CA KDR		49	f	✓	✓	✓
11	17-Aug-10	B	Fatty Liver	Fatty Liver			43	m	✓		✓
12	17-Nov-10	B	Fatty Liver	Fatty Liver			40	m	✓	✓	✓
13	13-Aug-08	B	Fibrosis+	Fibrosis, Primary sclerosing cholangitis	TP53		51	f	✓	✓	✓
14	4-Jan-09	B	Fibrosis+	Fibrosis or Autoimmune hepatitis-sarkoidosis			34	m	✓		✓
15	18-May-10	B	Fibrosis+	Fibrosis, Primary Sclerosing Cholangitis	PIK3CA KIT		30	f	✓	✓	✓
16	6-Aug-10	B	Fibrosis+, Alcohol	Alcoholic Fibrosis, NASH			31	m	✓	✓	
17	14-Jan-09	B	Cirrhosis+, Alcohol	Alcoholic Cirrhosis, NASH	TP53		47	m	✓	✓	✓
18	19-Apr-11	B	Fibrosis+, Alcohol, Obesity	Fatty Liver, Alcoholic Cirrhosis, Obesity	RET		51	m	✓	✓	✓
19	2-Apr-14	B	HCC, HCV, Diab2	HCC, Chronic Hepatitis C, Cirrhosis, Diabetes Type 2	TP53 PIK3CA KIT KDR		58	m	✓	✓	✓
20	2-Apr-14	B	HCC, Obesity	HCC, Cirrhosis, NASH, Obesity	TP53 KIT		70	m	✓	✓	✓
21	2-Nov-14	B	HCC, Diab1	HCC, Diabetes Type 1	TP53 KDR		71	m	✓		✓
22	25-Feb-14	B	HCC	HCC, Cirrhosis	TP53 KDR		72	m	✓		✓
23	3-Nov-14	B	HCC, Diab2, Alcohol	HCC, Alcoholic Cirrhosis, Diabetes Type 2	MET		63	m	✓	✓	✓
24	25-Mar-14	B	HCC, HCV, Diab2, Obesity	HCC, Cirrhosis, Chronic Hepatitis C, Diabetes Type 2, Obesity	TP53 TP53 TP53 PIK3CA CTNNB1 KDR	TP53 CTNNB1	69	m	✓		✓
25	4-Aug-14	B	HCC, HCV	HCC, Chronic Hepatitis C	KIT KDR		59	m	✓	✓	✓
26	24-Jul-14	B	HCC, Alcohol	HCC, Alcoholic Cirrhosis	TP53 KDR		65	m	✓	✓	✓
27	29-Jul-14	B	HCC, HAV, HBV	HCC, Chronic Hepatitis B and A, Cirrhosis	PTEN	PTEN	35	f	✓		✓
28	21-Aug-14	B	HCC, Diab2	HCC, Cirrhosis, NASH, Diabetes Type 2	TP53		65	f	✓	✓	✓
29	26-Aug-14	B	HCC	HCC, Cirrhosis	TP53 TP53 PIK3CA	TP53	73	m	✓	✓	✓
30	26-Aug-14	B	HCC, Alcohol	HCC, Alcoholic Cirrhosis			77	m	✓	✓	✓
31	28-Aug-14	B	HCC, HCV	HCC, Chronic Hepatitis C	TP53 KDR		73	f	✓	✓	✓
32	9-Feb-14	B	HCC	HCC			67	m	✓		
33	9-Apr-14	B	HCC	HCC			80	f	✓	✓	
34	23-Sep-14	B	HCC	HCC	TP53		78	m	✓	✓	✓
35	23-Sep-14	B	HCC	HCC	TP53 PIK3CA CTNNB1 CDKN2A	PIK3CA CTNNB1 CDKN2A	84	m	✓	✓	✓
36	27-May-09	B	HCV, Fatty Liver	Chronic Hepatitis C, <5% fatty liver			51	m	✓	✓	✓
37	23-Jun-09	B	HCV, Fatty Liver	Chronic Hepatitis C, 20% fatty liver	TP53 KDR		47	m	✓	✓	✓
38	12-Feb-09	B	HCV, Fatty Liver	Chronic Hepatitis C, very mild fatty liver	TP53 PIK3CA KDR	TP53	38	m	✓	✓	✓
39	15-Dec-10	B	HCV	Chronic Hepatitis C	KDR		60	m	✓		✓
40	2-Aug-11	B	HCV, Fatty Liver	Chronic Hepatitis C, 5% fatty liver	TP53	TP53	39	f	✓	✓	✓
41	16-Feb-11	B	HCV	Chronic Hepatitis C	KDR		31	m	✓	✓	✓
42	24-Nov-10	B	HCV, Fatty Liver	Chronic Hepatitis C, 5% fatty liver	TP53 KDR APC	APC	53	f	✓	✓	✓
43	1-Dec-10	B	HCV, Fatty Liver	Chronic Hepatitis C, 10% fatty liver	TP53 KDR SMARCB		40	f	✓		✓
44	22-Nov-06	B	HCV	Chronic Hepatitis C	TP53 KDR SMARCB MET		45	f	✓	✓	✓
45	29-Mar-11	B	HCV, Fatty Liver	Chronic Hepatitis C, 5% fatty liver	TP53		38	f	✓	✓	✓
46	2-Apr-09	B	Fatty Liver, Fibrosis, Obesity	Fatty Liver (40%), mild fibrosis, Obesity	TP53		46	m	✓		✓
47	4-Sep-08	B	Fatty Liver, Fibrosis	Fatty Liver, mild fibrosis			30	m	✓		
48	9-Dec-07	B	Fatty Liver, Fibrosis	Fatty Liver, mild fibrosis	SMARCB KDR		48	m	✓	✓	✓
49	3-Jul-07	B	Fatty Liver, Obesity	Fatty Liver, Obesity	TP53		51	m	✓	✓	✓
50	8-Aug-06	B	Fatty Liver	Fatty Liver	TP53 PIK3CA KDR		31	m	✓		✓
51	30-May-06	B	Fatty Liver	Fatty Liver (30%)			64	f	✓	✓	✓
52	11-Sep-10	B	Fatty Liver, Fibrosis	Fatty Liver, mild fibrosis	TP53 KDR		58	f	✓	✓	✓
53	2-Jun-08	B	Fatty Liver	Fatty Liver (60%)	TP53 APC	APC	55	m	✓	✓	✓
54	18-Jun-08	B	Fatty Liver, Fibrosis	Fatty Liver (70%), mild fibrosis	KDR MET		46	m	✓	✓	✓
55	6-Apr-08	B	Fatty Liver	Fatty Liver (40%), med-tox	TP53 KDR		36	m	✓	✓	✓
56	22-Sep-08	B	HCV, Fibrosis+, Obesity	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis, Obesity	TP53 KDR		57	m	✓		✓
57	23-Jul-08	B	HCV, Fibrosis+	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis, 60% fatty liver			44	m	✓	✓	✓
58	15-Jul-08	B	HCV, Fibrosis+	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis, 10% fatty liver	TP53	TP53	47	m	✓	✓	✓

59	20-Feb-08	B	HCV, Fibrosis+	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis	KDR		64	f	✓	✓	✓
60	12-Dec-07	B	HCV, Fibrosis+	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis	TP53 MET		51	m	✓	✓	✓
61	19-Jun-07	B	HCV, Fibrosis+	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis	TP53		65	f	✓	✓	✓
62	16-May-07	B	HCV, Fibrosis+	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis, 10% fatty liver	TP53 MET		59	f	✓	✓	✓
63	28-Mar-07	B	HCV, Fibrosis+	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis, 10% fatty liver			52	m	✓	✓	✓
64	4-Nov-07	B	HCV, Fibrosis+, Alcohol	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis, alcoholic fatty liver (30%)	KDR NOTCH1	NOTCH1	47	m	✓		✓
65	5-Sep-07	B	HCV, Fibrosis+, Obesity	Chronic Hepatitis C, Advanced Fibrosis/Cirrhosis, Obesity	TP53 MET	TP53	45	m	✓		✓
66	4-Jul-12	S	HCC, HCV	HCC, Hepatitis-C-Cirrhosis	KDR		67	m	✓	✓	✓
67	12-Aug-12	S	HCC	HCC, Child-Pugh A Cirrhosis	TP53 CTNNB1 KDR	CTNNB1	56	m	✓	✓	✓
68	24-Aug-12	S	HCC	HCC	KDR		73	m	✓	✓	✓
69	4-Sep-12	S	HCC	HCC, Fibrosis/Cirrhosis	TP53 KDR		57	m	✓	✓	✓
70	21-Aug-12	S	HCC	HCC	ATM TP53 SMARCB1 PIK3CA KDR	ATM	86	m	✓	✓	✓
71	16-Oct-12	S	HCC	HCC, Child-Pugh A Cirrhosis	KDR TP53	TP53	76	m	✓	✓	✓
72	28-Nov-12	S	HCC, HCV	HCC, Hepatitis C, Steatosis hepatis (ca. 80 %)	KDR		58	m	✓	✓	✓
73	19-Dec-12	S	HCC, Nicotine	HCC, Nicotine addiction	KDR MET		79	m	✓	✓	✓
74	7-Jan-13	S	HCC Recidive	Recidive HCC, Cirrhosis			83	m	✓		
75	5-Feb-13	S	HCC, HBV	HCC, Hepatitis B, Neoplastic vascular invasion, Fibrosis			40	f	✓	✓	✓
76	29-Apr-13	S	HCC	HCC, Cirrhosis	CTNNB1 TP53	CTNNB1 TP53	59	m	✓	✓	✓
77	19-Apr-13	S	HCC	HCC	TP53 KDR		61	f	✓	✓	✓
78	3-Sep-13	S	HCC	HCC			78	f	✓	✓	✓
79	7-Oct-13	S	HCC	HCC	TP53		67	f	✓	✓	✓
80	17-Oct-13	S	HCC, Alcohol, Nicotine	HCC, Neoplastic vascular invasion, Nicotine addiction, Alcohol addiction	TP53 KDR CDKN2A	TP53	62	m	✓	✓	✓
81	20-Dec-13	S	HCC, HBV	HCC, Hepatitis B	TP53 KDR	TP53	60	m	✓	✓	✓
82	7-Mar-14	S	HCC	HCC, Cirrhosis	SMARCB1 KIT KDR		73	m	✓	✓	✓
83	11-Mar-14	S	HCC	HCC	ATM CTNNB1 KDR	ATM CTNNB1	77	m	✓	✓	✓
84	12-Mar-14	S	HCC, Recidive	Recidive HCC	TP53		74	m	✓	✓	✓
85	2-Apr-14	S	HCC	HCC	KDR		72	m	✓	✓	✓
86	16-Apr-14	S	HCC	HCC	PTEN TP53 KDR	PTEN	62	f	✓	✓	✓
87	29-Apr-14	S	HCC	HCC, Child-Pugh A Cirrhosis	TP53		59	m	✓	✓	✓
88	6-May-14	S	HCC, HAV	HCC, Hepatitis A	KDR		76	m	✓	✓	✓
89	9-May-14	S	HCC	HCC, Child-Pugh A Cirrhosis	KDR		68	m	✓	✓	✓
90	25-Jul-14	S	HCC, HAV, HBV	HCC, Hepatitis B, Hepatitis A	TP53 PIK3CA	TP53	54	m	✓	✓	✓
91	28-Oct-14	S	HCC	HCC	TP53 STK11 CTNNB1 KDR	STK11 CTNNB1	53	m	✓	✓	✓
92	25-Nov-14	S	HCC, HBV	HCC, Hepatitis B	TP53 SMARCB1		70	f	✓	✓	✓
93	1-Dec-14	S	HCC, HBV, Alcohol	HCC, Hepatitis B, Alcohol addiction	PTEN TP53 KDR TP53	PTEN TP53	54	m	✓	✓	✓
94	9-Mar-15	S	HCC	HCC	TP53 KDR		55	m	✓	✓	✓
95	31-Mar-15	S	HCC	HCC, Steatosis hepatis	CTNNB1	CTNNB1	61		✓	✓	✓

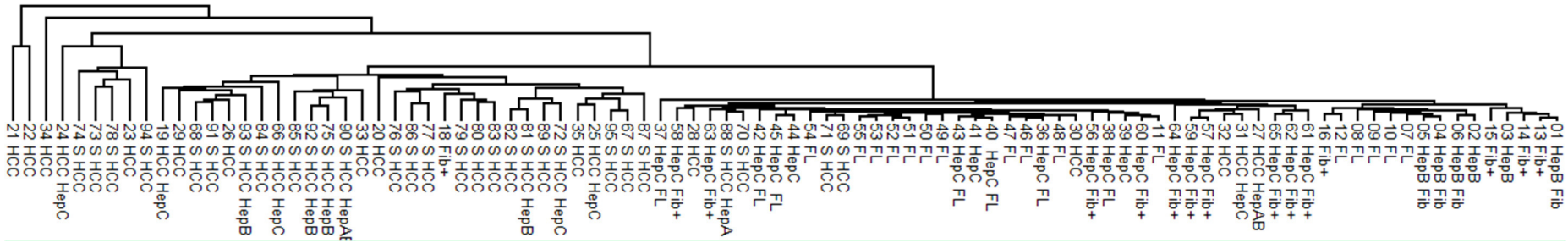
Term	**Meaning**
B	Biopsy
S	Surgery
HAV	Hepatitis A virus
HBV	Hepatitis B virus
HCV	Hepatitis C virus
Fibrosis+	Advanced Fibrosis/Cirrhosis
NASH	Non-alcoholic steatohepatitis
Diab1	Diabetes Type 1
Diab2	Diabetes Type 2
	Data not available
✓	Analysis performed

Appendix 4

Pearson



Spearman



Euclidian